DOCUMENT RESUME

ED 031 407 SE 007 159

By-Domotor, Zoltan
Probabilistic Relational Structures and Their Applications
Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.
Pub Date May 69
Note-171p.; Technical Report No. 144 Psychology Series
EDRS Price MF-$0.75 HC-$8.65
Descriptors-Logic, *Mathematical Applications, *Mathematical Concepts, *Mathematics, *Probability, *Probability Theory

The principal objects of the investigation reported were, first, to study qualitative probability relations on Boolean algebras, and secondly, to describe applications in the theories of probability logic, information, automata, and probabilistic measurement. The main contribution of this work is stated in 10 definitions and 20 theorems. The basic concern in this technical report was to show that probability, entropy, and information measures can be studied successfully in the spirt of representational or algebraic measurement theory. The method utilized in this report is based on the most general results of modern mathematics, which state a one-to-one correspondence among relations, cones in vector spaces, and the classes of positive functionals. (RP)

PROBABILISTIC RELATIONAL STRUCTURES AND THEIR APPLICATIONS

BY

ZOLTAN DOMOTOR

TECHNICAL REPORT NO. 144

May 14, 1969

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

TECHNICAL REPORTS

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

(Place of publication shown in parentheses; if published title is different from title of Technical Report, this is also shown in parentheses.)

(For reports no. 1 - 44, see Technical Report no. 125.)

50 R. C. Atkinson and R. C. Calfee. Mathematical learning theory. January 2, 1963. (In B. B. Wolman (Ed.), Scientific Psychology. New York: Basic Books, Inc., 1965. Pp. 254-275)

51 P. Suppes, E. Crothers, and R. Weir. Application of mathematical learning theory and linguistic analysis to vowel phoneme matching in Russian words. December 28, 1962.

52 R. C. Atkinson, R. Calfee, G. Sommer, W. Jeffrey and R. Shoemaker. A test of three models for stimulus compounding with children. January 29, 1963. (J. exp. Psychol., 1964, 67, 52-58)

53 E. Crothers. General Markov models for learning with inter-trial forgetting. April 8, 1963.

54 J. L. Myers and R. C. Atkinson. Choice behavior and reward structure. May 24, 1963. (Journal math. Psychol., 1964, 1, 170-203)

55 R. E. Robinson. A set-theoretical approach to empirical meaningfulness of measurement statements. June 10, 1963.

56 E. Crothers, R. Weir and P. Palmer. The role of transcription in the learning of the orthographic representations of Russian sounds. June 17,. 1963.

57 P. Suppes. Problems of optimization in learning a list of simple items. July 22, 1963. (In Maynard W. Shelly, II and Glenn L. Bryan (Eds.), Human Judgments and Optimality. New York: Wiley. 1964. Pp. 116-126)

58 R. C. Atkinson and E. J. Crothers. Theoretical note: all-or-none learning and intertrial forgetting. July 24, 1963.

59 R. C. Calfee. Long-term behavior of rats under probabilistic reinforcement schedules. October 1, 1963.

60 R. C. Atkinson and E. J. Crothers. Tests of acquisition and retention, axioms for paired-associate learning. October 25, 1963. (A comparison of paired-associate learning models having different acquisition and retention axioms, J. math. Psychol., 1964, 1, 285-315)

61 W. J. McGill and J. Gibbon. The general-gamma distribution and reaction times. November 20, 1963. (J. math. Psychol., 1965, 2, 1-18)

62 M. F. Norman. Incremental learning on random trials. December 9, 1963. (J. math. Psychol., 1964, 1, 336-351)

63 P. Suppes. The development of mathematical concepts in children. February 25, 1964. (On the behavioral foundations of mathematical concepts. Monographs of the Society for Research in Child Development, 1965, 30, 60-96)

64 P. Suppes. Mathematical concept formation in children. April 10, 1964. (Amer. Psychologist, 1966, 21, 139-150)

65 R. C. Calfee, R. C. Atkinson, and T. Shelton, Jr. Mathematical models for verbal learning. August 21, 1964. (In N. Wiener and J. P. Schoda (Eds.), Cybernetics of the Nervous System: Progress in Brain Research. Amsterdam, The Netherlands: Elsevier Publishing Co., 1965. Pp. 333-349)

66 L. Keller, M. Cole, C. J. Burke, and W. K. Estes. Paired associate learning with differential rewards. August 20, 1964. (Reward and information values of trial outcomes in paired associate learning. (Psychol. Monogr., 1965, 79, 1-21)

67 M. F. Norman. A probabilistic model for free-responding. December 14, 1964.

68 W. K. Estes and H. A. Taylor. Visual detection in relation to display size and redundancy of critical elements. January 25, 1965, Revised 7-1-65. (Perception and Psychophysics, 1966, 1, 9-16)

69 P. Suppes and J. Donio. Foundations of stimulus-sampling theory for continuous-time processes. February 9, 1965. (J. math. Psychol., 1967, 4, 202-225)

70 R. C. Atkinson and R. A. Kinchla. A learning model for forced-choice detection experiments. February 10, 1965. (Br. J. math stat. Psychol., 1965, 18, 184-206)

71 E. J. Crothers. Presentation orders for items from different categories. March 10, 1965.

72 P. Suppes, G. Groen, and M. Schlag-Rey. Some models for response latency in paired-associates learning. May 5, 1965. (J. math. Psychol., 1966, 3, 99-128)

73 M. V. Levine. The generalization function in the probability learning experiment. June 3, 1965.

74 D. Hansen and T. S. Rodgers. An exploration of psycholinguistic units in initial reading. July 6, 1965.

75 B. C. Arnold. A correlated urn-scheme for a continuum of responses. July 20, 1965.

76 C. Izawa and W. K. Estes. Reinforcement-test sequences in paired-associate learning. August 1, 1965. (Psychol. Reports, 1966, 18, 879-919)

77 S. L. Blehart. Pattern discrimination learning with Rhesus monkeys. September 1, 1965. (Psychol. Reports, 1966, 19, 311-324)

78 J. L. Phillips and R. C. Atkinson. The effects of display size on short-term memory. August 31, 1965.

79 R. C. Atkinson and R. M. Shiffrin. Mathematical models for memory and learning. September 20, 1965.

80 P. Suppes. The psychological foundations of mathematics. October 25, 1965. (Colloques Internationaux du Centre National de la Recherche Scientifique. Editions du Centre National de la Recherche Scientifique. Paris: 1967. Pp. 213-242)

81 P. Suppes. Computer-assisted instruction in the schools: potentialities, problems, prospects. October 29, 1965.

82 R. A. Kinchla, J. Townsend, J. Yellott, Jr., and R. C. Atkinson. Influence of correlated visual cues on auditory signal detection. November 2, 1965. (Perception and Psychophysics, 1966, 1, 67-73)

83 P. Suppes, M. Jerman, and G. Groen. Arithmetic drills and review on a computer-based teletype. November 5, 1965. (Arithmetic Teacher, April 1966, 303-309.

84 P. Suppes and L. Hyman. Concept learning with non-verbal geometrical stimuli. November 15, 1968.

85 P. Holland. A variation on the minimum chi-square test. (J. math. Psychol., 1967, 3, 377-413).

86 P. Suppes. Accelerated program in elementary-school mathematics -- the second year. November 22, 1965. (Psychology in the Schools, 1966, 3, 294-307)

87 P. Lorenzen and F. Binford. Logic as a dialogical game. November 29, 1965.

88 L. Keller, W. J. Thomson, J. R. Tweedy, and R. C. Atkinson. The effects of reinforcement interval on the acquisition of paired-associate responses. December 10, 1965. (J. exp. Psychol., 1967, 73, 268-277)

89 J. I. Yellott, Jr. Some effects on noncontingent success in human probability learning. December 15, 1965.

90 P. Suppes and G. Groen. Some counting models for first-grade performance data on simple addition facts. January 14, 1966. (In J. M. Scandura (Ed.), Research in Mathematics Education. Washington, D. C.: NCTM, 1967. Pp. 35-43.

(Continued on inside back cover)

PROBABILISTIC RELATIONAL STRUCTURES AND THEIR APPLICATIONS

by

Zoltan Domotor

TECHNICAL REPORT NO. 144

May 14, 1969

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

ACKNOWLEDGMENTS

I wish to express my sincere thanks to Professor Patrick Suppes and Professor Dana Scott for suggesting the problems, and for supervising the research leading to this dissertation. Their unfailing interest in the progress of my research, and their counsel, have played a large part in bringing this work to a successful conclusion. Many hours with Professor Suppes provided encouragement as well as valuable guidance and support. The knowledge and understanding I have gained in discussions with Professor Scott are invaluable to me.

Consultations with Professors Jaakko Hintikka and Andrzej Ehrenfeucht are gratefully appreciated.

I wish to express my thanks also to Dr. Juraj Bolf from the Institute of Measurement Theory of the Czechoslovak Academy of Sciences whose personal support and encouragement brought me to Stanford.

I am much indebted to Mr. David Miller for correcting the language of this work.

Finally, the major financial support received from the Institute for Mathematical Studies in the Social Sciences at Stanford University during my academic program at Stanford is also acknowledged with gratitude.

TABLE OF CONTENTS

CHAPTER

1. INTRODUCTION

1.1. Statement of Problems

The principal objects of the investigation reported here are, first, to study qualitative probability relations on Boolean algebras, and secondly, to describe applications in the theories of probability logic, information, automata, and probabilistic measurement.

Several authors (for example, B. de Finetti, B. O. Koopman, L. J. Savage, D. Scott, P. Suppes) have posed the following specific problems:

(P_1) Given a Boolean algebra $\mathcal{A}$ and a binary relation $\precsim$ on $\mathcal{A}$, under what conditions on $\precsim$ does there exist a probability measure P on $\mathcal{A}$ satisfying

$$A \precsim B \Longleftrightarrow P(A) \leq P(B)$$

for all $A, B \in \mathcal{A}$?

(P_2) Given a Boolean algebra $\mathcal{A}$ and a binary relation $\succ$ on $\mathcal{A}$, under what conditions on $\succ$ does there exist a probability measure P on $\mathcal{A}$ and a real number $0 < \varepsilon < 1$ satisfying

$$A \succ B \Longleftrightarrow P(A) \geq P(B) + \varepsilon$$

for all $A, B \in \mathcal{A}$?

(P_3) Given a Boolean algebra $\mathcal{A}$ and a quaternary relation $\precsim$ on $\mathcal{A}$, under what conditions on $\precsim$ does there exist a conditional probability measure P on $\mathcal{A}$ satisfying

$$A/B \precsim C/D \Longleftrightarrow P(A/B) \leq P(C/D)$$

for all $A, B, C, D \in \mathcal{A}$ for which $P(B) \cdot P(D) > 0$?

(P_4) Given a Boolean algebra $\mathcal{A}$ and a quaternary relation $\succ$ on $\mathcal{A}$, under what conditions on $\succ$ does there exist a conditional probability measure P on $\mathcal{A}$ and a real number $0 < \varepsilon < 1$ satisfying

$$A/B \succ C/D \Longleftrightarrow P(A/B) \geq P(C/D) + \varepsilon$$

for all $A, B, C, D \in \mathcal{A}$ for which $P(B) \cdot P(D) > 0$?

Chapter 2 answers problems (P_2), (P_3), and (P_4). The axioms for entropy originally given by Shannon in 1948 have been replaced several times subsequently by weaker conditions. In each case the axiomatization of the basic information-theoretic notions is presented as a collection of <u>functional equations</u>. In contrast, a new approach is proposed here; an approach which is an application of the techniques developed in the study of probabilistic relational structures. We shall give axiomatic definitions of the concepts of <u>qualitative information</u> and <u>qualitative entropy structure</u>; and we shall study some of their basic measurement-theoretic properties. For this purpose we also set down axiomatization for the <u>qualitative probabilistic independence</u> relations on both the algebra of <u>events</u> and the algebra of <u>experiments</u>.

Many methodologists have in recent years been leaning towards the view that as long as there is no satisfactory theory of the probability theory of first-order formulas, the rather delicate questions of inductive logic, confirmation theory and scientific method are not likely to be satisfactorily answered. Here it is argued that, if there is any truth in this view, a purely qualitative treatment of the probabilities of quantified formulas is a more promising line of attack than the quantitative theories propagated by Carnap and others.

In the mathematical theory of probability conditional probabilities are conditional probabilities of events of the basic algebra; in no sense are they probabilities of conditional events. But it seems an interesting problem whether they could be constructed in this second way. A definition of an algebra of such conditional events is given here which conforms to the intuitive concepts used by probability. Once we have a qualitative theory of probability, it is natural to ask if we can treat qualitatively all problems formulated in terms of a probability space. The algebraic character of probabilistic automata makes this a promising field of application, and in this work definitions of qualitative probabilistic automata are suggested. As further applications several empirical structures relevant in physics and social sciences are studied.

The investigation has produced many new problems in this field, and the main ones are listed in the conclusion.

1.2. Previous Results

There are several ways of introducing the concept of probability. In all of them, throughout the long history of the subject, the intention has been to answer the following two basic questions:

(Q_1) What are the entities, called events, which are supposed to be probable?

(Q_2) What kind of function or relation, called probability, is attributed to the events?

The main answers are usually referred to as measure-theoretical (H. Steinhaus [1], A. N. Kolmogorov [23]), limiting frequency (R. von Mises [3], A. Wald [4]), subjectivist (B. de Finetti [5], L. J. Savage [6]), logical (R. Carnap [7], H. Jeffreys [8]) and finally, methodological (R. B. Braithwaite [9]). Motivations for some of these answers to questions (Q_1) and (Q_2) are hidden in the complex problem of rationality.

The answer to question (Q_1) is algebraic: the set of events, structurally speaking, forms at least a lattice, and almost always a Boolean algebra, or, equivalently, a field of sets. There is less agreement on whether the events themselves should be interpreted as sets, statements, or perhaps sets of statements. But there is no obvious reason why all these should not be possible.

Question (Q_2) causes real trouble. In fact, this question is just what the foundations of probability are all about.

In this work we shall restrict ourselves to the study of the relationships between the formal structures of the measure-theoretical and subjectivist approaches.

De Finetti's subjectivist probability theory is written in terms of a binary relation $\precsim$, defined on some Boolean algebra $\mathcal{A}$ of events. The intended interpretation of $\precsim$, called the qualitative probability relation, is as follows:

> If $A, B \in \mathcal{A}$, then $A \precsim B$ means that the event A is (a priori) not more probable than the event B.

It is useful to define $A \prec B$ as $\neg\, B \precsim A$, and $A \sim B$ as $A \precsim B \;\&\; B \precsim A$.

The celebrated axioms of de Finetti's probability theory impose certain constraints on the qualitative probability relation, in order to guarantee the existence of a numerical probability measure on $\mathcal{A}$ in the standard sense; this problem was called (P_1) in Section 1.1. It turned out that de Finetti's conditions were necessary, but not sufficient; (P_1) was finally solved for the finite case by C. H. Kraft, J. W. Pratt, and A. Seidenberg in 1959 [10]. A more simple general solution was found by Scott in 1964 (D. Scott [11]). Scott has also obtained a solution for infinite Boolean algebras (D. Scott [12]).

The intended interpretation of the relation $\succ$ in problem (P_2) is as follows:

> $A \succ B \iff$ the event A is definitely more probable than event B $(A, B \in \mathcal{A})$.

Obviously $\succ$ is intended to be a semiordering relation; we shall call it a semiordered qualitative probability relation.

Problem (P_2) was raised by Suppes, and for finite Boolean algebras was first considered by J. H. Stelzer in his doctoral dissertation (J. H. Stelzer [13]), where a partial solution was given. The solution is deficient in that the necessary and sufficient conditions are not stated purely in terms of the qualitative relation $\succ$ (see Stelzer [13], Theorem 3.14, p. 68); moreover, the proof of the main theorem (ibid., Theorem 3.8, p. 52) is invalid.

B. O. Koopman [14], A. Shimony [15], and more recently P. Suppes [16] and R. D. Luce, investigated a more complicated case, considering conditional events. Well known is Koopman's relatively strong and complicated system of axioms for the binary relation $\preccurlyeq$, which is interpreted as follows:

> $A/B \preccurlyeq C/D \iff$ the event A, given event B is not more probable than the event C, given event D, where $A, B, C, D \in \mathcal{A}$.

For criticism, applications to confirmation theory, and a further review of this problem, we refer to Shimony [15]. We should perhaps mention here that Koopman's approach has the following defects. It contains axioms like $A/B \preccurlyeq C/D \Longrightarrow (B \subseteq A \Longrightarrow D \subseteq C)$, so that the qualitative probability relation imposes certain Boolean relations on the events; This is implausible if $\preccurlyeq$ is not connected, that is,

$$A/B \preccurlyeq C/D \vee C/D \preccurlyeq A/B ,$$

which for some reason is the only case Koopman is prepared to consider.

However, one of his axioms pretty well amounts to postulating the existence of equi-probable partitions of arbitrary events, which is impossible in non-trivial finite cases.

By far the best system of axioms known to the author for the relation $\preccurlyeq$, the <u>qualitative conditional probability relation</u>, was given by Suppes [16]. Unfortunately, his axioms are necessary, but not sufficient. This is obvious, since they are first-order axioms; and even in the case of (P_1) a second-order axiom is needed. Besides that, without sufficient conditions we have no way of representing one probability structure by another.

Problem (P_4) was first discussed in Suppes [16] (in connection with the problems of causality), where necessary conditions are given for the relation $\succ$, the <u>semiordered qualitative conditional probability relation</u>. The intended interpretation is obvious: $A/B \succ C/D$ means that event A given event B is definitely more probable than event C given event D.

As far as the author knows, no solutions to the problems (P_2), (P_3), and (P_4) have yet been given.

We would like to emphasize that we shall primarily be interested in the cases where the Boolean algebra $\mathcal{A}$ is finite. For atomless Boolean algebras, for instance, it is quite easily shown that, under certain rather natural conditions on $\preccurlyeq$, there is only one probability measure compatible with $\preccurlyeq$ in the sense of problem (P_1). Such a result for σ-algebras was given by C. Villegas [17] as a generalization of certain investigations of L. J. Savage.

In probability theory, or rather in its foundations, there has long been a trend towards identifying events with formulas of certain first-order formalized languages. Among principal proponents of this idea we can certainly count J. M. Keynes, H. Jeffreys, H. Reichenbach, R. Carnap, and J. Łukasiewicz.

It is of course formally possible to ascribe probability to formulas, since, under rather simple conditions, they form a Boolean algebra. Yet a perfect solution to this problem for (quantified) formulas is not as simple as this makes it sound.

For example, if we investigate the theory of <u>linear ordering structures</u>, $\mathcal{M} = \langle M, \leq \rangle$, we can ask for the probability of the formula $x \leq y$ for $x, y \in M$. If we say, for instance, that $P(x \leq y) = 1/2$, then this should mean in the frequency interpretation that by drawing in a given way the elements x, y from M, we obtain pairs which in one half of the cases will satisfy the formula $x \leq y$. But, although $P(x \leq y)$ may equal $1/2$, nevertheless $P(\forall_x \forall_y (x \leq y))$ can hardly be anything but 0; for this universal sentence is false in any non-trivial ordered set. How about the probability of $\exists_x \forall_y (x \leq y)$? It depends, of course, on the structure $\mathcal{M}$ in question. If $\mathcal{M}$ is a suitable structure, then the formula will be true or false in it, and hence will have probability 1 or 0.

A theory that can only attribute probabilities of 0 or 1 to sentences is inadequate for almost all applications. But alternative approaches may lead to more satisfactory probability assignments. One way is to assume that we are given a set of possible worlds from which one world can be chosen at random. In this world we perform another random drawing, this time of elements of the world. Then the probability of a formula is equal to the probability of its being satisfied by the double drawing. More technically, we first draw a model $\mathcal{M}$ in accordance with a given probability measure ν on the family $\mathbb{M}$ of all models under consideration; and then from $\mathcal{M}$, again in accordance with a probability measure $\mu_{\mathcal{M}}$ given in $\mathcal{M}$, we draw a set of elements.

Every formula Φ has a probability $\mu_{\mathcal{M}}(\Phi)$ in the selected model $\mathcal{M}$. Keeping Φ constant, and allowing the model $\mathcal{M}$ to vary, we obtain a random variable $\mu_{\mathcal{M}}(\Phi)$, for which we can compute the expected value $\mathbb{E}\,\mu_{\mathcal{M}}(\Phi)$ with respect to the probability measure ν, defined on the family $\mathbb{M}$. Hence, the probability of the formula Φ is given by

$$P(\Phi) = \int_{\mathbb{M}} \mu_{\mathcal{M}}(\mathcal{M}[\Phi])\, d\,\nu\,, \quad \text{where}$$

$$\mathcal{M}[\Phi] = \{v\colon \Phi \text{ is true in } \mathcal{M} \text{ under valuation } v\}\,.$$

In the case of conditional probabilities, the conditional expectation would do the job. These ideas are due to J. Łoś [18].

Gaifman [19] also developed a theory of probabilities on formulas of arbitrary first-order languages, and proved that a rather natural way of extending to quantified formulas a probability measure defined on molecular formulas was in fact unique. Scott and Krauss [20] then generalized Gaifman's method to infinitary languages. Ryll-Nardzewski realized that assigning probabilities to formulas is just a special case of the well-known method of assigning values in complete Boolean algebras.

It should be pointed out that, whatever its other merits, probability logic by no means exhausts the problems in probability theory. On the contrary, nearly all the methods and results of the mathematical theory, especially those involving random variables, expectations, and limits, far outstrip probability logic. Nevertheless, as mentioned above, there are many interesting results, several of them peculiar to this field.

The author's aim will be to survey these developments from the point of view of qualitative probability theory, and to apply them to probabilistic measurement theory.

Automata theory, as a part of abstract algebra, is a well-developed discipline, whereas probabilistic automata theory is still in a more or less primitive state. The most important work on this problem is due to M. O. Rabin and D. Scott [21] and P. H. Starke [22]; and qualitative versions of some of their definitions will be given in Chapter 4.

1.3. Contribution of this Research

Most of the contributions have already been described; here they are briefly summarized.

The central mathematical results are the solutions of (P_2), (P_3), and (P_4).

The author proposes a new interpretation of the conditional event A/B. Systematic axiomatic development of conditional probability theory has been done by A. Rényi [23, 24] and A. Császár [25].

In the present author's opinion, the answer to (Q_2) for the conditional case cannot be satisfactorily answered if question (Q_1) for conditional events is not already answered.

Using the proof technique of problems (P_2) - (P_4) the author succeeded in obtaining several representation theorems for information and entropy structures. In connection with these structures considerable attention has been devoted to the qualitative independence relations on events and on experiments.

In the final chapter certain results of probability logic are handled anew by qualitative methods. Qualitative probabilistic notions are also applied to probabilistic automata theory and probabilistic measurement structures.

1.4. Methodological Remarks

One of the more fruitful ways of analyzing the mathematical structure of any concept is what we here call the representation method.

This method consists of determining the entire family of homomorphisms or isomorphisms from the analyzed structure into a suitable well-known concrete mathematical structure. The work is usually done in two steps: first, the existence of at least one homomorphism is proved; secondly, one finds a set or group of transformations up to which the given homomorphism is exactly specified. The unknown and analyzed structure is then represented by a better known and more familiar structure, so that eventually, the unknown problem can be reduced to one perhaps already solved.

Another advantage of this method is that it handles problems of empirical "meaning" and content in an extensional way. For it is a rather trivial fact that any mathematical approach to such a problem will give the answer at most up to isomorphism. Hence all meaning problems are extramathematical questions. For example, interpretation of the concept of probability is beyond the scope of the Kolmogorov axioms.

Yet, without permanently flying off on a tangent, we would like to indicate by an example (anyway needed in the sequel) how by using the idea of representation of one structure by another one can handle the "meaning" problem inside mathematics.

The next two chapters will deal with certain mathematical structures. The problems these structures pose are too difficult to answer immediately, and we shall therefore translate the problem into geometric language by means of the representation of relations by cones in a vector space. From this geometric language we translate again into functional language, by means of the representation of cones by positive functionals. Here the problem is solved, and we translate the result back into the original language of relations.

This is one of the most efficient ways of thinking in mathematics. It should be noted, however, that the translation is not always reversible. The representing structure may keep only one aspect of the original structure, but this has the advantage that the problem may be stripped of inessential features, and replaced by a familiar type of problem, hopefully easier to solve. Of course, essential features may be lost. In spite of this, the method of sequential representation has proved its worth in a great variety of successful applications.

Take as a concrete example the relational structure of the qualitative probability $< \mathcal{A}, \precsim >$ which will be discussed extensively in Chapter 2; any empirical content assigned to the probability structure $< \mathcal{A}, \precsim >$ is carried through the chain of homomorphisms: relational entity $\rightsquigarrow$ geometric entity $\rightsquigarrow$ functional entity, to the probability measure P on $\mathcal{A}$. The measure P may thus acquire empirical content on the basis of the structure $< \mathcal{A}, \precsim >$ which we assume already to have empirical content via other structures or directly, by stipulation.

In general, the empirical meaning or content of an abstract, or so-called *theoretical* structure (model) is given through a more or less complicated tree or lattice of structures together with their mutual homomorphisms (satisfying certain conditions), where some of them, the initial, concrete, or so-called *observational* ones, are endowed with empirical meanings by postulates.

Note that the homomorphism is here always a *special function*. For example, in the case of probability, P satisfies *not only* the *homomorphism condition* (which is relatively simple), but also the *axioms* for the probability measure. Thus the axioms for the given structure are *essentially involved* in the existence of the homomorphism. In this respect, the representation method goes far beyond the ordinary homomorphism technique between similar structures, or the theory of elementarily equivalent models.

The "meaning" of a given concept can be expressed extensionally by the lattice of possible representation structures connected mutually by homomorphisms (with additional properties) and representing always one particular aspect of the concept.

We do not intend to go into this rather intricate philosophical subject here. The only point of this discussion was to emphasize the methodological importance of our approach to concepts like qualitative probability, information and entropy.

2. QUALITATIVE PROBABILITY STRUCTURES

2.1. Algebra of Events

We start with some prerequisites for answering the question (Q_1) in Section 1.1. Probability theory studies the mathematical properties of the structure $< \mathcal{B}, Q >$, where $\mathcal{B}$ is a Boolean algebra and Q is a probability measure on $\mathcal{B}$; or the structure $\mathbb{A} = < \Omega, \mathcal{E}, P >$, where Ω is a nonempty set of <u>sample points</u>, $\mathcal{E}$ is a field of subsets of Ω, called the <u>field of events</u>, and P is a <u>probability measure</u> on $\mathcal{E}$.

These two structures, $< \mathcal{B}, Q >$ and $\mathbb{A}$ are closely related by the Stone's Representation Theorem, which says that every Boolean algebra $\mathcal{B}$ is isomorphic to a field of sets $\mathcal{E}_S$, that is, $\mathcal{B} \cong \mathcal{E}_S \; (= \mathcal{E})$.

Those authors who work with the structure $< \mathcal{B}, Q >$ do so largely because no commitment is made on the character of the elements of a Boolean algebra (it does not really matter whether they are sets or propositions or something else); a further advantage is that one can treat the probability as a strictly positive measure, and forget about the events of measure zero, which have no probabilistic meaning anyway. On the other hand, the concept of a random variable can hardly be defined in this structure in a direct way. So for applications the second structure, $\mathbb{A}$, is more convenient. An interesting attempt to reduce the notion of a random variable to that of a σ-homomorphism of a field of Borel sets of real numbers into a Boolean σ-algebra was made by R. Sikorski [26]. Though this succeeds, nothing more general is gained by it, as thus it really matters little which structure we take as our primary object of study.

There are good reasons to keep both structures in mind; one is that there is a probabilistic interpretation of the Stone isomorphism between the Boolean algebras $\mathcal{S}$ and $\mathcal{U}$.

In particular, if we start with the model $<\mathcal{S}, Q>$, and if $\mathcal{S} \cong \mathcal{U}_S$ as above, then (see Halmos [27]) $\mathcal{U}_S$ is the field of closed-and-open (clopen) sets of a zero-dimensional (or totally disconnected) compact Hansdorff space Ω_S which is associated with the family of all prime ideals of $\mathcal{S}$, and therefore also ultrafilters of $\mathcal{S}$.

Without loss of generality, we can think of Ω_S as the set of ultrafilters of $\mathcal{S}$. On Ω_S we then can define random variables in the standard way, so that from $<\mathcal{S}, Q>$ we can get $<\Omega_S, \mathcal{U}_S, P_S>$ by adopting the measure Q into P_S by isomorphism. The converse should be obvious.

By analogy with mathematical logic, where the collection of all formulas of a formalized first-order language is, roughly speaking, identified with a Boolean algebra, a theory is identified with a filter, a complete theory with an ultrafilter, and so on, we shall provide similar, probabilistic identifications.

For this purpose let $\mathbb{A}$ be a standard probability space as described above.

In current textbooks of probability theory it is customary to consider the notion of the occurrence of an event as a monadic primitive predicate Θ.

If ΘA means that event A occurs, then it is rather trivial to check that the following formulas are valid for all $A, B \in \mathcal{A}$:

(i) $\Theta\, \Omega$,

(ii) $A \subseteq B \ \&\ \Theta A \Longrightarrow \Theta B$,

(iii) $\Theta A \ \&\ \Theta B \Longrightarrow \Theta A \cap B$,

(iv) $\Theta A \vee \Theta \bar{A}$.

Set-theoretically this means that the set of all events occurring at a given trial forms an ultrafilter: $\nabla = \{A\colon \Theta A \ \&\ A \in \mathcal{A}\}$.

Naturally $\Delta = \{\bar{A}\colon A \in \nabla\}$ is a maximal ideal, so that the set of events which do not occur at a given trial forms a maximal (or prime) ideal: $\Delta = \{A\colon \neg\, \Theta A \ \&\ A \in \mathcal{A}\}$. But then $\mathcal{A} = \Delta \cup \nabla$; that is to say, each trial (or experiment) decomposes the algebra of events $\mathcal{A}$ into two disjoint structures Δ and ∇ .

If we call the outcome of a trial that element ω of Ω which is the true result of the trial, then the principal ultrafilter ∇

is generated by the singleton $\{\omega\}$, so that we should write $\nabla(\{\omega\})$ instead of ∇. Similarly, the prime ideal Δ is generated by $\overline{\{\omega\}}$, so that we shall write $\Delta(\overline{\{\omega\}})$ instead of Δ. Therefore, any trial can be viewed as an ordered couple $< \nabla(\{\omega\}), \Delta(\overline{\{\omega\}}) >$, where ω is the outcome of the trial.

Summarizing, we conclude that:

$\nabla(\{\omega\})$ = the set of those events which occur at the outcome ω of the given trial.

$\Delta(\overline{\{\omega\}})$ = the set of those events which do not occur at the outcome ω of the given trial.

Let $\nabla(A)$ be the filter generated by A; then since

$$\nabla(A) = \bigcap_{\omega \in A} \nabla(\{\omega\}),$$

$\nabla(A)$ = the set of those events which occur in all outcomes $\omega \epsilon A$.

Similarly, since $\Delta(A) = \bigcap_{\omega \in \bar{A}} \Delta(\overline{\{\omega\}})$,

$\Delta(A)$ = the set of those events which do not occur in any of the outcomes $\omega \epsilon \bar{A}$.

Especially,

$$\nabla(\Omega) = \{\Omega\} \quad \text{and} \quad \Delta(\emptyset) = \{\emptyset\}, \quad \text{hence}$$

the only event which occurs at all possible outcomes is Ω, and the only event which fails to occur at any outcome is $\emptyset$.

The set of all principal ideals $\mathcal{J} = \{\Delta(A) : A \in \mathcal{U}\}$ is isomorphic to $\mathcal{U}$: $\mathcal{J} \cong \mathcal{U}$, if we define in $\mathcal{J}$ the Boolean opeartions as follows:

$$\triangle(A) + \triangle(B) = \triangle(A \cup B) ,$$
$$\triangle(A) \cdot \triangle(B) = \triangle(A \cap B) ,$$
$$\overline{\triangle(A)} = \triangle(\bar{A}) .$$

Using the analytic properties of the sequences of ultrafilters, we can give a rigorous definition of the frequency-interpretation of probability.

The isomorphism φ , constructed by Stone, has also a probabilistic interpretation. If $A \in \mathcal{S}$, then $\varphi(A) = \{ \nabla : A \in \nabla \, \& \, \nabla \in \Omega_S \}$, where, as pointed out before, Ω_S is the set of all ultrafilters of $\mathcal{S}$. Hence, $\varphi(A)$ is nothing else but the set of <u>all</u> <u>experiments</u> (trials) in which A <u>occurs</u>. Obviously $\varphi(\Omega) = \Omega_S$ and $\varphi(\emptyset) = \emptyset$.

Having this interpretation in mind, we shall freely use in the sequel both the structures $< \mathcal{S}, Q >$ and $\mathbb{A} = < \Omega, \mathcal{E}\mathcal{X}, P >$.

Next we shall characterize set-theoretically the notion of a <u>conditional</u> <u>event</u>. Remember that in probability theory one speaks only about the <u>conditional</u> <u>probability</u> <u>of</u> <u>an</u> <u>event</u> $(P_B(A))$ and such a thing as the <u>probability</u> <u>of</u> <u>a</u> <u>conditional</u> <u>event</u> $(P(A/B))$ does not exist, since the entity, <u>conditional</u> <u>event</u>, is not defined.

On the other hand, applied probability is full of interpretations of conditional probabilities which encourage us to believe in the existence of conditional events as independent entities.

The present study needs conditional events for several purposes; rather than postulate their existence, we honestly set about giving them a satisfactory set-theoretic definition.

From the one-one correspondence between filters $\mathcal{F}$ and ideals $\mathcal{J}$ we obtain an isomorphism $\mathcal{F} \cong \mathcal{J}$, where the atoms of $\mathcal{F}$ are the ultrafilters $\nabla(\{\omega\})$, $\omega\epsilon\Omega$.

Using the isomorphism between the lattice of ideals of $\mathcal{O}\!\mathcal{L}$ and the lattice of congruence relations on $\mathcal{O}\!\mathcal{L}$, we can introduce the following equivalence relation on $\mathcal{O}\!\mathcal{L}$:

$$A \equiv B \bmod \Delta \iff A \uplus B \in \Delta \quad (A, B \in \mathcal{O}\!\mathcal{L}); {}^{*)}$$

$$\Delta = \{A : A \equiv \emptyset \;\&\; A \in \mathcal{O}\!\mathcal{L}\}.$$

By duality, we get the congruence relation also for filters:

$$A \equiv B \bmod \nabla \iff A \leftrightarrow B \in \nabla \qquad (A, B \in \mathcal{O}\!\mathcal{L}); {}^{**)}$$

$$\nabla = \{A : A \equiv \Omega \;\&\; A \in \mathcal{O}\!\mathcal{L}\}.$$

In particular,

$$A \equiv B \bmod \nabla(C) \iff A \equiv B \bmod \Delta(\overline{C}) \iff AC = BC.$$

The probabilistic interpretation of the congruence relation $\equiv$ is the following:

$A \equiv B \bmod \nabla(\{\omega\}) \iff$ the events A and B are <u>indistinguishable</u>, given the outcome ω. More generally, $A \equiv B \bmod \nabla(C) \iff$ the events A and B are <u>indistinguishable</u>, given all the outcomes in C; that is, $\Theta A \iff \Theta B$, given $\omega\epsilon C$.

We can introduce this indistinguishability relation $\equiv$ into the algebra of events $\mathcal{O}\!\mathcal{L}$ by constructing quotient Boolean algebras $\mathcal{O}\!\mathcal{L}/\nabla(A)$ or $\mathcal{O}\!\mathcal{L}/\Delta(\overline{A})$.

The reader will notice that

$$\mathcal{O}\!\mathcal{L}/\Delta(\overline{\emptyset}) = \mathcal{O}\!\mathcal{L}/\Delta(\Omega) \cong \{\mathcal{O}\!\mathcal{L}\} \cong \mathcal{O}\!\mathcal{L}/\nabla(\emptyset).$$

---

*) $A \uplus B$ denotes symmetric difference, that is, $A \uplus B = A\overline{B} \cup \overline{A}B$.

**) $A \leftrightarrow B$ denotes $\overline{A}\overline{B} \cup AB$.

Therefore we shall rule out this pathological Boolean algebra by putting $A \neq \emptyset$.

On the other hand, the ultrafilters $\nabla(\{\omega\})$ and maximal ideals $\Delta(\overline{\{\omega\}})$ generate two-element quotient Boolean algebras: $\mathcal{Cl}/\nabla(\{\omega\}) \cong \{ \mathbb{0}, \mathbb{1} \} \cong \mathcal{Cl}/\Delta(\overline{\{\omega\}})$ where $\mathbb{1}$ corresponds to $\nabla(\{\omega\})$, and $\mathbb{0}$ to $\Delta(\overline{\{\omega\}})$; that is, $[\Omega]_{\equiv} = \nabla(\{\omega\})$ and $[\emptyset]_{\equiv} = \Delta(\overline{\{\omega\}})$.

We have given plenty of examples that show that it does not matter whether we consider ideals or filters. Filters are more convenient for conventional thinkers; we think in terms of <u>occurred events</u>, rather than the non-occurred ones. From now on, therefore, we shall work only in terms of filters.

If we put $\mathcal{Cl}/A = \mathcal{Cl}/\nabla(A) \quad (A \neq \emptyset)$, then $\mathcal{Cl}/A$ can be interpreted as the Boolean algebra of <u>conditional</u> <u>events</u>, conditionalized by event A. Hence for any $B \in \mathcal{Cl}$, B/A is a conditional event, equal to the class of events, <u>indistinguishable</u> from event B, given the outcomes in A.

By considering $\mathcal{Cl}/A \quad (A \neq \emptyset)$ we restrict the set of possible outcomes to the set A. Naturally, $\mathcal{Cl}/\Omega \cong \mathcal{Cl}$, so that conditionalization by Ω is trivial. The conditional event B/A takes care also of the fact that the probability of the event B depends only on the intersection of B and A. Thus, if $B/A = C/A$, then $AB = AC$, which is obviously true. If $B \cap \mathcal{Cl} = \{B \cap A : A \in \mathcal{Cl}\}$, for $B \in \mathcal{Cl}$, then it is easy to check that the following isomorphisms are valid:

$\triangle(B) \cong \mathfrak{A}/\triangle(\bar{B}) \cong B \cap \mathfrak{A} \cong \mathfrak{A}/\nabla(B)$. Hence, the study of $\mathfrak{A}/B$ is the study of the same probability structure as before, but with the set of possible outcomes restricted.

Now naturally, in order to define a suitable measure P^* in $\mathfrak{A}/B$, given the probability space $\mathbb{A}$, we have to realize that the conditional event A/B is a <u>sure event</u> if and only if it always occurs, that is, if $A \in \nabla(B)$. Moreover, since $P^*(A/B) = P^*(C/B)$ if $A/B = C/B$, we must have $P^*(A/B) = P^*(C/B)$ if $P(AB) = P(CB)$. Due to the fact, pointed out before, that $P^*(\Omega/B) = P^*(A/B) = 1$, if $A \in \nabla(B)$, we are bound to accept $P^*(A/B)$ as simply $\frac{P(A \cap B)}{P(B)}$ $(P(B) > 0)$.

To sum up, if we are given a probability space $\mathbb{A}$, then any <u>restriction</u> of the set of possible outcomes leads to conditionalization and therefore to an appropriate conditional measure.

It is clear how to interpret the following Boolean operations in the set of conditional events $\mathfrak{A}/B$:

$$A/B + C/B = A \cup C/B\ ,$$
$$A/B \cdot C/B = A \cap C/B\ ,$$
$$\overline{A/B} = \bar{A}/B\ .$$

Similarly, the meaning of the identities $A/B = AB/B$, $AB/BC = A/BC$ should be clear enough.

The reader may wonder where the multiplicative law for conditional probabilities is hidden. It can be checked that

$$\mathfrak{A}/B \cap C \cong (\mathfrak{A}/C)\Big/B/C\ , \qquad (2.1)$$

which means that we can assign isomorphically $A/C\Big/B/C = AB/C\Big/B/C$

to $A/B \cap C$. We can also check that the measure in $(\mathfrak{A}/C)/B/C$ should be $\frac{P^*(AB/C)}{P^*(B/C)}$ (once we are given the measure P^* in $\mathfrak{A}/C$) and that this measure is just the same as $P^{**}(A/B \cap C)$ in $\mathfrak{A}/B \cap C$; hence

$$P^*(A/B \cap C) \cdot P^*(B/C) = P^*(A \cap B/C) , \qquad (2.2)$$

if the appropriate algebraic existence conditions are satisfied. We proceed analogously as in the case of P^*.

Note that (2.1) together with (2.2) state a simple fact, namely, that iterated restriction of the domain of possible outcomes Ω by B, and then C, amounts to the simultaneous restriction of Ω by B and C; that is, its restriction by $B \cap C$. (2.2) is the most important relation between two conditional events, conditionalized differently. To take the set $\{A/B : A, B \in \mathfrak{A}, B \neq \emptyset\}$ and look for some structure in it is not reasonable; for what can we expect to get in the set $\bigcup_{A \in \mathfrak{A}} \mathfrak{A}/A$, which is not even a lattice? To take the direct sum $\bigoplus_{A \in \mathfrak{A}} \mathfrak{A}/A$ is much more reasonable. We shall reserve a place for discussion of this algebraic construction in Section 2.6 on qualitative conditional probabilities. Our main concern in this section was to give set-theoretic definitions of the notions of occurrence, trial, and conditional event, and to explain the main relationships between probability measures on Boolean algebras and fields of events. An interesting notion of conditional probability is presented in H. P. Evans and S. C. Kleene [28]; on the other hand, a radical attempt to uncover some structure in the set of conditional entities can be found in A. H. Copeland [29].

2.2. Basic Facts about Qualitative Probability Structures

In this section we discuss the results of Scott [11, 12] concerning the problem (P_1). The general method applied here goes back to a theorem of Mazur and Orlicz [30] (see p. 174, Theorem 2.41). This theorem is a simple generalization of the well-known Hahn-Banach theorem on the extention of linear functionals in normed linear spaces. Mazur and Orlicz's Theorem gives in rather good terms a necessary and sufficient condition for the solvability of a system of linear inequalities. As an application one could hope to solve problems related to ours, provided that they involve showing the existence of a linear functional on a given set which would homomorphically match a relation defined on this set. Because of its importance, we shall presently quote the generalized version of this theorem.

Before we proceed to the details on the relation between <u>ordered structures</u> and <u>linear functionals</u> we shall recall briefly a couple of notions from the theory of ordered vector spaces needed in the sequel.

A real vector space equipped with an ordering compatible with its linear structure is called an <u>ordered vector space</u>. More specifically, given a real vector space $\mathcal{V}$ and a binary relation $\precsim$ on $\mathcal{V}$, then the couple $< \mathcal{V}, \precsim >$ is called an <u>ordered real vector space</u> if and only if

(i) $\precsim$ is reflexive, transitive, connected, and antisymmetric;

(ii) $\forall v_1, v_2, w \in \mathcal{V} \; [v_1 \precsim v_2 \Longrightarrow v_1 + w \precsim v_2 + w]$;

(iii) $\forall v_1, v_2, \alpha \in Re^{+} [v_1 \precsim v_2 \rightarrow \alpha v_1 \precsim \alpha v_2]$. *)

An equivalent definition can be given in terms of a cone. By a positive cone in a vector space, we mean a nonempty subset $\mathcal{C} \subseteq \mathcal{V}$ such that the following geometric properties are satisfied for all $v, w \in \mathcal{V}$:

(a) $v, w \in \mathcal{C} \Rightarrow v + w \in \mathcal{C}$;

(b) $\alpha \in Re^{+} \;\&\; v \in \mathcal{C} \Rightarrow \alpha v \in \mathcal{C}$;

(c) $v \in \mathcal{C} \;\&\; -v \in \mathcal{C} \Rightarrow v = 0$;

(d) $v \in \mathcal{C} \vee -v \in \mathcal{C}$.

The link between the ordering relation $\precsim$ and the cone $\mathcal{C}$ is given by

$$v \precsim w \Leftrightarrow w - v \in \mathcal{C} \qquad \text{for all } v, w \in \mathcal{V} .$$

Hence, the notion of an ordered vector space can be given equivalently in terms of the structure $< \mathcal{V}, \mathcal{C} >$. The reader will remember our discussion of the translation of relation-theoretic notions into geometric ones in Section 1.4. If $\mathcal{C}$ satisfies only (a) and (b), it is called a wedge. Hence, in particular, a wedge is a convex subset of $\mathcal{V}$. If the ordering $\precsim$ in $\mathcal{V}$ allows us to construct a supremum and infimum for each subset of $\mathcal{V}$, then $< \mathcal{V}, \precsim >$ is called a lattice-ordered vector space.

---

*) Re^{+} denotes the set of non-negative real numbers.

Since there is a close relationship between the ordering and the topological structure of the ordered vector space, this is reflected in the specific nature of linear mappings on these spaces. Thus we can translate the geometric notions into functional ones, as pointed out in Section 1.4. This fact is hidden in the Mazur and Orlicz generalization of the classical Hahn-Banach Theorem.

THEOREM 1 (Mazur-Orlicz) *Let* $\mathcal{V}$ *be a vector space and* $\langle \mathcal{W}, \preccurlyeq \rangle$ *a complete lattice-ordered vector space. If* $\varphi : \mathcal{V} \longrightarrow \mathcal{W}$ *is a mapping for which*

$$\varphi(v + w) \preccurlyeq \varphi(v) + \varphi(w) \quad \text{for all } v, w \in \mathcal{V} ;$$

$$\varphi(\alpha\, v) = \alpha\, \varphi(v) \quad \text{for all } \alpha \in Re^+, v \in \mathcal{V} ,$$

and if $\{v_i\}_{i\in I} \subseteq \mathcal{V}$ & $\{w_i\}_{i\in I} \subseteq \mathcal{W}$, *then there is a linear mapping* $\Phi : \mathcal{V} \longrightarrow \mathcal{W}$ *such that*

(i) $w_i \preccurlyeq \Phi(v_i)$ *for all* $i \in I$,

(ii) $\Phi(v) \preccurlyeq \varphi(v)$ *for all* $v \in \mathcal{V}$,

if and only if

for any $\{i_1, i_2, \ldots i_n\} \subseteq I$ *and* $\{\alpha_1, \alpha_2, \ldots \alpha_n\} \subseteq Re^+$ *it is true that*

$$\sum_{k=1}^{n} \alpha_k w_{i_k} \preccurlyeq \varphi\left(\sum_{k=1}^{n} \alpha_k v_{i_k}\right) .$$

There are several known proofs of this theorem. We shall use the argument of V. Pták [31].

The necessity of the inequality is clear since Φ is a linear mapping.

To show the sufficiency, suppose that the inequality holds. For given $v \in \mathcal{V}$, and $\{i_1, \ldots, i_n\} \subseteq I$ & $\{\alpha_1, \alpha_2, \ldots \alpha_n\} \subseteq Re^+$ the following is true:

$$\sum_{k=1}^{n} \alpha_k w_{i_k} \preccurlyeq \varphi\left(v + \sum_{k=1}^{n} \alpha_k v_{i_k}\right) + \varphi(-v) ;$$

thus

$$-\varphi(-v) \preccurlyeq \varphi\left(v + \sum_{k=1}^{n} \alpha_k v_{i_k}\right) - \sum_{k=1}^{n} \alpha_k w_{i_k} .$$

Since $\langle \mathcal{W}, \preccurlyeq \rangle$ is a complete lattice-ordered vector space, we can define $\psi : \mathcal{V} \longrightarrow \mathcal{W}$ by

$$\psi(v) = \inf_{\substack{i_k \in I,\ k \leq n \\ \alpha_k \in Re^+}} \left\{ \varphi\left(v + \sum_{k=1}^{n} \alpha_k v_{i_k}\right) - \sum_{k=1}^{n} \alpha_k v_{i_k} \right\} .$$

We can show very fast that $\psi(\alpha v) = \alpha\psi(v)$ for $v \in \mathcal{V}$ & $\alpha \in Re^+$. In addition, if $i_k \in I$, $\alpha_i \in Re^+$, $i \leq n$, and $j_k \in I$, $\beta_k \in Re^+$, $k \leq m$, and $v_1, v_2 \in \mathcal{V}$, then

$$\varphi\left(v_1 + \sum_{k=1}^{n} \alpha_k w_{i_k}\right) - \sum_{i=1}^{n} \alpha_k w_{i_k} + \varphi\left(v_2 + \sum_{k=1}^{m} \beta_k w_{j_k}\right) - \sum_{i=1}^{m} \beta_k w_{j_k} \succcurlyeq$$

$$\succcurlyeq \varphi\left(v_1 + v_2 + \sum_{k=1}^{n} \alpha_k w_{i_k} + \sum_{k=1}^{m} \beta_k w_{j_k}\right) - \sum_{k=1}^{n} \alpha_k w_{i_k} - \sum_{k=1}^{m} \beta_k w_{j_k} \succcurlyeq$$

$$\succcurlyeq \psi(v_1 + v_2) .$$

Hence, $\psi(v_1 + v_2) \preccurlyeq \psi(v_1) + \psi(v_2)$ for $v_1, v_2 \in \mathcal{V}$.

Using the assumption of the theorem, we can derive the existence of $\Phi : \mathcal{V} \longrightarrow \mathcal{W}$ with the required properties. Q.E.D.

The following corollary is a simple consequence of the previous theorem:

COROLLARY 1 <u>If</u> $< \mathcal{V}, \| \| >$ <u>is a normed vector space and</u> $\delta : \mathcal{V} \longrightarrow \mathrm{Re}$ <u>is a functional on</u> $\mathcal{V}$, <u>then there is a linear functional</u> $\varphi : \mathcal{V} \longrightarrow \mathrm{Re}$ <u>such that</u> $\delta(v) \leq \varphi(v) \leq \|v\|$ <u>for all</u> $v \in \mathcal{V}$

if and only if

$$\sum_{k \leq m} \delta(v_k) \leq \| \sum_{k \leq m} v_k \|$$

<u>holds for all</u> $\{v_k\}_{k \leq m} \subseteq \mathcal{V}$.

This corollary can be used to find out what kind of conditions should be imposed on a wedge $\mathcal{C}$ in $\mathcal{V}$ in order to guarantee the existence of a linear functional φ on $\mathcal{V}$, and positive on $\mathcal{C}$.

For any wedge $\mathcal{C}$ of $\mathcal{V}$ we set $\mathcal{C}^+ = \mathcal{C} - \{-v : v \in \mathcal{C}\}$; that is to say, we remove from $\mathcal{C}$ those vectors whose negative counterparts are also in $\mathcal{C}$.

For $\mathcal{S} \subseteq \mathcal{V}$ we define $C^+[\mathcal{S}]$ as the set $\{\sum_{i \leq m} \alpha_i v_i : v_i \in \mathcal{S}, \alpha_i > 0, i \leq m\}$, and call it a <u>positive linear closure</u> of $\mathcal{S}$. We set $\| \mathcal{S} \| = \inf_{v \in \mathcal{S}} \{\|v\|\}$, if $\mathcal{S} \subseteq \mathcal{V}$.

Now we are ready to state a theorem proved by Scott [12]:

THEOREM 2 (Representation Theorem) <u>Let</u> $\mathcal{C} \subseteq \mathcal{V}$ <u>be a wedge of the normed vector space</u> $< \mathcal{V}, \| \| >$. <u>Then the necessary and sufficient</u>

condition for the existence of a linear functional $\varphi : \mathcal{V} \longrightarrow Re$ such that for all $v \in \mathcal{V}$:

(i) $\varphi(v) \leq \| v \|$;

(ii) $v \in \mathcal{C} \Longrightarrow \varphi(v) \geq 0$;

(iii) $v \in \mathcal{C}^+ \Longrightarrow \varphi(v) > 0$;

is the following:

$$\exists \{k_i\}_{i=1}^{\infty} \Big[\ k_i \text{ is convex } \& \ k_i \subseteq \mathcal{V} \ \&$$

$$\& \ \mathcal{C}^+ = \mathcal{C}^+ [\bigcup_{i=1}^{\infty} k_i] \ \& \ \| k_i + \mathcal{C} \| > 0 \text{ for}$$

$$\text{all } i = 1, 2, \ldots \Big] .$$

Proof:

I. Let $\varphi : \mathcal{V} \longrightarrow Re$ satisfy (i) - (iii). Then $\mathcal{C}^+ = \{v \in \mathcal{C}: \varphi(v) > 0\}$. If we define $k_i = \{v \in \mathcal{C} : \varphi(v) \geq 1/i\}$ for $i = 1, 2, \ldots,$ then k_i is convex and $\mathcal{C}^+ = \mathcal{C}^+[\bigcup_{i=1}^{\infty} k_i]$. (i) & (ii) imply:

$v \in k_i$ & $w \in \mathcal{C} \Longrightarrow 1/i \leq \varphi(v) + \varphi(w) \leq \| v + w \|$, thus

$1/i \leq \| k_i + \mathcal{C} \|$.

II. Let $\{k_i\}_{i=1}^{\infty}$ be a sequence of convex sets in $\mathcal{V}$ satisfying the conclusion of the theorem. Let us define $\delta_i : \mathcal{V} \longrightarrow Re$ for $i = 1, 2, \ldots$ as follows:

$$\delta_i(v) = \begin{cases} \|k_i + \mathcal{C}\|, & \text{if } v \in k_i ; \\ 0 , & \text{if } v \in \mathcal{C} - k_i ; \\ -\infty , & \text{otherwise.} \end{cases}$$

If $v_k \in k_i$ for $k \leq m$ and $w \in \mathcal{E}$, then

$$\| k_i + \mathcal{C} \| \leq \| 1/m \cdot \sum_{k \leq m} v_k + 1/m \cdot w \| \text{, because}$$

k_i is convex. Clearly δ_i satisfies the conclusion of Corollary 1.

Let us define linear functionals $\varphi_i : \mathcal{V} \longrightarrow \mathrm{Re}$ such that $\delta_i(v) \leq \varphi_i(v) \leq \| v \|$ $(v \in \mathcal{V})$, according to Corollary 1, and put

$$\varphi(v) = \sum_{i=1}^{\infty} \frac{\varphi_i(v)}{2i} .$$

Then φ satisfies (i). $\varphi_i(v) \geq 0$ for $v \in \mathcal{C}$ implies (ii). By virtue of the definition, δ_i: $(v) \geq \|k_i + \mathcal{C}\| > 0$ for $v \in k_i$;

thus $\varphi(v) > 0$, for $v \in \bigcup_{i=1}^{\infty} k_i$, so that (iii) also is true for φ.

Q. E. D.

COROLLARY 2 <u>Let $\mathcal{C} \subseteq \mathcal{V}$ be a wedge of the normed vector space $< \mathcal{V}, \| \, \| >$ with countable basis</u> B, <u>that is,</u> $\mathcal{C} = C^+[B] \cup \{0\}$. <u>Then</u> $v \in \mathcal{C}^+ \implies \| v + \mathcal{C} \| > 0$ <u>is the necessary and sufficient condition for the existence of a linear functional</u> $\varphi : \mathcal{V} \longrightarrow \mathrm{Re}$ <u>satisfying</u> (i) - (iii) <u>of Theorem</u> 2.

Proof: Put $T = B \cap \mathcal{C}^+$; then $\mathcal{C}^+ = C^+[\bigcup_{v \in T} (v + \mathcal{C})]$, since if $w = \sum_{k \leq m} \alpha_k v_k \in \mathcal{C}^+$, where $v_k \in B$ and $\alpha_k > 0$ for $k \leq m$, then $\exists\, k_0 \leq m\, [v_{k_0} \in \mathcal{C}^+]$. Thus

$$w = \alpha_{k_0}(v_{k_0} + \sum_{\substack{k \leq m \\ k \neq k_0}} \frac{\alpha_k}{\alpha_{k_0}} v_k), \text{ which means that } w \in C^+[v_{k_0} + \mathcal{C}].$$

If the basis B of the wedge $\mathcal{E}$ in Corollary 2 is finite, then

$$\| v + \mathcal{E} \| > 0 \Longleftrightarrow v \notin -\mathcal{E} .$$

But it is always true that $v \in \mathcal{E}^{+} \Longrightarrow v \notin -\mathcal{E}$. Thus in the finite dimensional case, the functional φ always exists.

Theorem 2 has basic importance. We can translate binary relations on $\mathcal{V}$ into cones in $\mathcal{V}$ as explained earlier, and then show the existence of a linear functional on $\mathcal{V}$ satisfying certain monotony conditions.

As an important consequence, we shall prove, using Scott's unpublished notes, the following theorem:

THEOREM 3 <u>Let</u> $< \mathcal{U}, \precsim >$ <u>be a structure, where</u> $\mathcal{U}$ <u>is a Boolean algebra with zero element</u> $\emptyset$ <u>and unit element</u> Ω , <u>and</u> $\precsim$ <u>is a binary relation on</u> $\mathcal{U}$ <u>such that</u>

$$\emptyset \prec \Omega, \quad \emptyset \precsim A, \quad \text{and}$$

$$A \precsim B \vee B \precsim A \quad \text{for all} \quad A, B \in \mathcal{U} .$$

<u>Further, let</u>

$$\mathcal{E}(\precsim) = \{ \sum_{i \leq m} \alpha_i(\hat{B}_i - \hat{A}_i) : A_i \precsim B_i \ \& \ \alpha_i \in Re^{+} \ \&$$

$$\& \ A_i, B_i \in \mathcal{U} \text{ for } i \leq m \} , \quad \text{where}$$

$\hat{A}$ <u>denotes a vector in the normed vector space of all continuous functions on the Stone space</u> Ω_S <u>of</u> $\mathcal{U}$ <u>with the usual supremum norm.</u>

<u>Then for there to exist a probability measure</u> P <u>on</u> $\mathcal{U}$ <u>such that</u>

$$A \precsim B \Longleftrightarrow P(A) \leq P(B) \quad \text{for all} \quad A, B \in \mathcal{U} ,$$

it is necessary and sufficient that there exist relations $\prec_i$ *on* $\mathcal{U}$, $i = 1, 2, \ldots$, *such that, for all* $A, B \in \mathcal{U}$,

(1) $A \prec B \iff A \prec_i B$ *for some* i,

(2) $\forall_{n,m} [1/n \leq \| 1/m \sum_{k \leq m} (\hat{A}_k - \hat{B}_k) + \mathcal{C}(\precsim) \|]$, *if*

$B_k \prec_n A_k$ *for* $k \leq m$ *and* $A_k, B_k \in \mathcal{U}$, $k \leq m$.

Proof:

I. Put $A \prec_i B \iff P(B) - P(A) \geq 1/i$, $i = 1, 2, \ldots$.

II. $\mathcal{C}^+(\precsim) = C^+[\bigcup \{\hat{A} - \hat{B} + \mathcal{C}(\precsim) : B \precsim A\}]$. Thus, if $k_n \subseteq \mathcal{V}$ is the convex set, generated by the set $\bigcup \{\hat{A} - \hat{B} + \mathcal{C}(\precsim) : B \precsim A\}$, then the conclusion of Theorem 2 is verified, as is easily seen. Therefore we obtain a linear functional $\varphi : \mathcal{V} \longrightarrow Re$ such that $\varphi(v) \leq \|v\|$ for $v \in \mathcal{V}$, and

$$\hat{A} \precsim \hat{B} \implies \varphi(A) \leq \varphi(B),$$

$$\hat{A} \prec \hat{B} \implies \varphi(A) < \varphi(B), \text{ if } A, B \in \mathcal{U}.$$

Since $\varphi(\hat{\Omega}) > 0$, $\varphi(\hat{A}) \geq 0$ for $A \in \mathcal{U}$, we can put $P(\hat{A}) = \frac{\varphi(\hat{A})}{\varphi(\hat{\Omega})}$, and, in view of $\hat{A} \precsim \hat{B} \iff A \precsim B$, also $P(A) = P(\hat{A})$. Q. E. D.

Remarks:

(1) The technique of identifying elements of a family of sets with vectors in a vector space $\mathcal{V}$ will be used over and over again. In

our case assigned one-one to the element $A \in \mathcal{A}$ is the <u>characteristic function</u> of the corresponding closed-and-open subset of the Stone space Ω_S of $\mathcal{A}$. This characteristic function, which is in the vector space $\mathcal{V}(\Omega_S)$, generated by the set Ω_S, will be denoted throughout the paper by $\hat{A}$. (See the discussion of the Stone space in 2.1.) In particular, if $A, B \in \mathcal{A}$, then $\hat{A} + \hat{B}$ is the sum in $\mathcal{V}(\Omega_S)$, and is equal to $(A \cup B)^\wedge$, provided $A \cap B = \emptyset$. If $\hat{\mathcal{A}} = \{\hat{A} : A \in \mathcal{A}\}$, then clearly $\hat{\mathcal{A}} \subseteq \mathcal{V}(\Omega_S)$.

(2) We shall keep in mind the well-known fact that each finitely additive probability measure on $\mathcal{A}$ is the restriction to $\hat{\mathcal{A}}$ of a <u>unique</u> linear functional $\varphi : \mathcal{V}(\Omega_S) \longrightarrow Re$ with $\varphi(v) \leq \|v\|$ for $v \in \mathcal{V}(\Omega_S)$ and $\varphi(\Omega) = 1$; and that the restriction of every such functional is a measure.

(3) Theorems like

$$A \precsim B \ \&\ B \precsim C \Longrightarrow A \precsim C ;$$

$$A \precsim B \ \&\ C \precsim D \Longrightarrow A \cup C \precsim B \cup D , \quad \text{if } A \perp C,\ B \perp D^{*)}$$

are easy consequences of the rather complicated conditions (1) and (2) of Theorem 3.

COROLLARY 3 <u>Let</u> $< \mathcal{A}, \precsim >$ <u>be a structure, where</u> $\mathcal{A}$ <u>is a countable Boolean algebra with zero element</u> $\emptyset$ <u>and unit element</u> Ω ; <u>let</u> $\precsim$ <u>be a binary relation on</u> $\mathcal{A}$ <u>such that</u> $\emptyset \prec \Omega$, $\emptyset \precsim A$, <u>and</u> $A \precsim B \vee B \precsim A$ <u>for all</u> $A, B \in \mathcal{A}$.

*) $A \perp B$ means $A \cap B = \emptyset$.

Then using the notation of Theorem 3, the necessary and sufficient condition for the existence of a probability measure P on $\mathcal{A}$ such that

$$A \precsim B \Longleftrightarrow P(A) \le P(B) \quad \text{for all} \quad A, B \in \mathcal{A}$$

is

$$\| \hat{A} - \hat{B} + \mathcal{C}(\precsim) \| > 0 , \quad \text{if} \quad B \prec A , \quad (A, B \in \mathcal{A}) .$$

Proof follows from Corollary 2.

If the Boolean algebra $\mathcal{A}$ is finite, then the condition in Corollary 3 boils down to a rather simple one, namely,

$$\forall_{i < n} [A_i \precsim B_i] \Longrightarrow B_n \precsim A_n , \quad \text{if} \quad \sum_{i=1}^{n} \hat{A}_i = \sum_{i=1}^{n} \hat{B}_i ,$$

where A_i, $B_i \in \mathcal{A}$ for $i \le m$.

COROLLARY 4 Let $< \Omega, \mathcal{A}, \precsim >$ be a structure, where Ω is a nonempty finite set; $\mathcal{A}$ is the Boolean algebra of subsets of $\mathcal{A}$ and $\precsim$ is a binary relation on $\mathcal{A}$.

Then the necessary and sufficient conditions for the existence of a probability measure P such that $< \Omega, \mathcal{A}, P >$ is a finitely additive probability space and

$$A \precsim B \Longleftrightarrow P(A) \le P(B) \quad \text{for all} \quad A, B \in \mathcal{A}$$

are the following:

(i) $\emptyset \prec \Omega$,

(ii) $\emptyset \precsim A$,

(iii) $A \precsim B \vee B \precsim A$,

(iv) $\forall_{i < n} [A_i \precsim B_i] \Longrightarrow B_n \precsim A_n$, if

$$\dot{\bigcup}_{i \leq n} A_i = \dot{\bigcup}_{i \leq n} B_i \ \& \ \dot{\bigcup}_{\substack{i,j \leq n \\ i < j}} A_i \cap A_j = \dot{\bigcup}_{\substack{i,j, \leq n \\ i < j}} B_i \cap B_j \ \& \ \ldots \ \&$$

$$\& \ A_1 \cap A_2 \cap \ldots \cap A_n = B_1 \cap B_2 \cap \ldots \cap B_n ,$$

where $A, B, A_i, B_i \in \mathcal{O}\!\mathcal{L}$ for $i = 1, 2, \ldots, n$ and $\dot{\bigcup}_{i \leq n} A_i$ denotes the symmetric difference of the n sets $A_1, A_2, \ldots, A_n$. (For two sets A_1 and A_2, $A_1 \dot{\cup} A_2$ is of course $(A_1 \cap \overline{A}_2) \cup (\overline{A}_1 \cap A_2)$.)

Proof:

First of all, the system of identities of symmetric differences of sets in (iv) is equivalent to $\sum_{i \leq n} \hat{A}_i = \sum_{i \leq n} \hat{B}_i$, where $\hat{A}$ denotes the characteristic function of the set A (Here the Stone space of $\mathcal{O}\!\mathcal{L}$ is identified with Ω.). Secondly, (iv) is equivalent to

$$A \prec B \longrightarrow \|\hat{B} - \hat{A} + \mathcal{C}(\precsim)\| > 0 \qquad (A, B \in \mathcal{O}\!\mathcal{L}).$$

For (a) assume (iv) and that

$A \precsim B$, but that

$\|\hat{B} - \hat{A} + \mathcal{C}(\precsim)\| = 0.$ Then

$\hat{A} - \hat{B} = \sum_{k \leq m} \alpha_k(\hat{A}_k - \hat{B}_k)$ for some m and for some

$\alpha_k \geq 0, \quad B_k \precsim A_k, \quad k \leq m.$

Since the characteristic function $\hat{A} - \hat{B}$ is integer-valued, we may assume that the scalars α_k are at least rational. By clearing fractions, transposing, and allowing repetitions, we may even assume that $\alpha_k = 1$ for all $k \leq m$. Hence

$$\sum_{k \leq m} \hat{A}_k + \hat{B} = \sum_{k \leq m} \hat{B}_k + \hat{A} .$$

But since $\forall_{k \leq m} B_k \precsim A_k$, by (iv) we get $B \precsim A$ which contradicts $A \prec B$.

(b) Clearly (iv) follows from $\|\hat{B} - \hat{A} + \mathcal{E}(\precsim)\| > 0$, if $A \prec B$.

Another easy consequence of Theorem 2 in finite case is the following corollary:

COROLLARY 5 Let $\mathcal{V}$ be a finite-dimensional real vector space and let $< M, \precsim >$ be a finite binary relational structure, where $\emptyset \neq M \subseteq \mathcal{V}$ and M is a set of vectors with rational coordinates with respect to some fixed basis of $\mathcal{V}$. Then there exists a linear functional $\varphi : \mathcal{V} \longrightarrow \mathrm{Re}$ such that for all $v, w \in M$

$$v \precsim w \iff \varphi(v) \leq \varphi(w)$$

if and only if

(1) $v \leq w \vee w \leq v$,

(2) $\forall_{i < n} [v_i \precsim w_i] \Longrightarrow w_n \precsim v_n$, if $\sum_{i \leq n} v_i = \sum_{i \leq n} w_i$,

where $v, w, v_i, w_i \in \mathcal{V}$ for $i = 1, 2, \ldots n$.

As we have seen, the conditions to be imposed on the Boolean algebra $\mathcal{Ol}$ enriched by a binary relation $\precsim$ in order to get a probability measure solving the problem (P_1) are rather simple in the finite case. On the other hand, the infinite case is utterly unintuitive. There may be some hope for simplifying the conditions in the infinite case, too, but we shall not deal with this problem here.

It is worth noting that Theorem 2 is general enough to be used in proving various representation theorems, important in algebraic measurement theory.

The structure $< \Omega, \mathcal{Ol}, \precsim >$, satisfying the conditions (i) - (iv), in Corollary 4, will be called a finite qualitative probability structure (FQP-structure). This notion can also be defined in terms of a strict ordering relation $\prec$, in which case, the axioms for $< \Omega, \mathcal{Ol}, \prec >$ to be a FQP-structure, are as follows:

(i) $\emptyset \prec \Omega$;

(ii) $\neg A \prec \emptyset$;

(iii) $A \prec B \Longrightarrow \neg B \prec A$;

(iv) $\forall_{i < n} [A_i \prec B_i] \Longrightarrow B_n \prec A_n$,

where $A, B, A_i, B_i \in \mathcal{Ol}$ for $1 \leq i \leq n$ and

$\sum_{i \leq n} \hat{A}_i = \sum_{i \leq n} \hat{B}_i$.

If we put

$$A \sim B \Longleftrightarrow (\neg A \prec B \ \& \ \neg B \prec A) ,$$

$$A \precsim B \Longleftrightarrow (A \prec B \vee A \sim B) ,$$

for all $A, B \in \mathcal{U}$, then the above definition becomes equivalent to Scott's definition, spelled out in Corollary 4; simply because $A \prec B \Longleftrightarrow \neg B \precsim A$. We shall freely use both definitions.

2.3. Additively Semiordered Qualitative Probability Structures

In Section 2.2 we discussed the general framework for the solution of problem (P_1), and we pointed out that the method used was general enough to be applied to other similar problems. The main task of this section will be to give the solution to problem (P_2).

The notion of a semiorder comes up when a set $\mathcal{U}$ is being ordered by some relation $\succ$ and, it is not always known whether two elements from $\mathcal{U}$ are indifferent. More precisely, a couple $\langle \mathcal{U}, \succ \rangle$ is called a semiorder structure iff*) it satisfies the following conditions for all $A, B, C, D \in \mathcal{U}$:

(i) $\neg A \succ A$;

(ii) $A \succ B \ \& \ C \succ D \Longrightarrow A \succ D \vee C \succ B$;

(iii) $A \succ B \ \& \ B \succ C \Longrightarrow A \succ D \vee D \succ C$.

The concept of a semiorder is due to R. D. Luce [32], and the axioms (i) - (iii) were given by Scott & Suppes [33].

---

*) iff is short for if and only if

If a semiorder structure $< \mathcal{O}\!\mathcal{L}, \succ >$ satisfies also

$$\text{(iv)} \quad A \succ B \Longrightarrow A \cup C \succ B \cup C \,, \quad \text{if } A, B \perp C \,, \; ^{*)}$$

then we shall call it an <u>additive semiorder structure</u>.

In this section we shall deal with <u>finite</u> additive semiorder structures $< \mathcal{O}\!\mathcal{L}, \succ >$; the <u>interpretation</u> of the formula $A \succ B$ for $A, B \in \mathcal{O}\!\mathcal{L}$ will be: event A is definitely more probable than event B.

(We prefer to use the symbol $\succ$ instead of $\prec$ because of the possible confusion with the <u>strict</u> qualitative probability relation discussed in the previous section.)

We assume the motivation for a semiorder relation $\succ$ to be known. Perhaps we should point out that semiorder is an adequate notion for representing algebraic measurement problems, in which the given measurement method has limited sensitivity, so that '<u>locally</u>' the <u>transitivity</u> for $\succ$ does not hold. In psychology one talks about the so-called <u>just noticeable difference</u> (<u>jnd</u>), whose appropriate numerical measure is a <u>fixed positive real number</u> $\mathcal{E}$ (which can be normalized to 1 by choosing a suitable unit). Hence $\mathcal{E}$ is a measure of the <u>threshold</u> of the measurement method.

For more sophisticated measurement problems we have to assume that <u>jnd</u> is <u>not constant</u>, but <u>varies</u> from one measured entity to another. For this purpose, Luce [32] introduced the notions of <u>lower</u> and <u>upper jnd</u> measures $\underline{\mathcal{E}}$ and $\overline{\mathcal{E}}$ which, in fact, define a <u>jnd interval</u>

---

*) $A \perp B$ means $A \cap B = \emptyset$. The other notation from set theory and logic is standard.

about each possible result of measurement.

Bearing all this in mind, we turn now to problem (P_2). For methodological reasons we prefer to start with the following definition:

DEFINITION 1 A triple $< \Omega, \mathcal{A}, \succ >$ is said to be a finitely additive semiordered qualitative probability structure (FASQP-structure) if and only if the following axioms are satisfied:

S_0 Ω is a nonempty finite set; $\mathcal{A}$ is the Boolean algebra of subsets of Ω ; and $\succ$ is a binary relation on $\mathcal{A}$;

S_1 $\Omega \succ \emptyset$;

S_2 $\neg A \succ A$;

S_3 $C \succ B \Longrightarrow C \succ A$, if $A \subseteq B$;

S_4 $\forall_{i < n} [A_i \succ B_i \ \& \ \neg C_i \succ D_i] \Longrightarrow [A_n \succ B_n \Longrightarrow C_n \succ D_n]$,

where $A, B, C, A_i, B_i, C_i, D_i \in \mathcal{A}$ for $1 \leq i \leq n$; and

$\sum_{i \leq n} (\hat{A}_i + \hat{D}_i) = \sum_{i \leq n} (\hat{B}_i + \hat{C}_i)$. ($\hat{A}$ denotes the characteristic function of the set A.)

Remarks:

(a) As pointed out before, the formula in axiom S_4 that concerns characteristic functions can easily be translated into a system of identities among sets, by means of the following fact:

$$\sum_{i \leq n} \hat{A}_i = \sum_{i \leq m} \hat{B}_i \quad \text{for } m \leq n \text{ , if and only if}$$

$$\dot{\bigcup_{i \leq n}} A_i = \dot{\bigcup_{i \leq m}} B_i \text{ ;}$$

$$\dot{\bigcup_{\substack{i,j \leq n \\ i < j}}} A_i A_j = \dot{\bigcup_{\substack{i,j \leq m \\ i < j}}} B_i B_j \text{ ;}$$

$$\dot{\bigcup_{\substack{i_1, i_2, \ldots, i_m \leq n \\ i_1 < i_2 < \ldots < i_m}}} A_{i_1} A_{i_2} \ldots A_{i_m} = B_1 B_2 \ldots B_m \text{ ;}$$

$$\dot{\bigcup_{\substack{i_1, i_2, \ldots, i_k \leq n \\ i_1 < i_2 < \ldots < i_k}}} A_{i_1} A_{i_2} \ldots A_{i_k} = \emptyset \text{ , if } m < k \leq n \text{ ,}$$

$$A_i, B_j \in \mathcal{Ol} \text{ , } 1 \leq i \leq n \text{ , } 1 \leq j \leq m \text{ .}$$

Thus the FASQP-structure is given by axioms which contain as primitives <u>only</u> the relation $\succ$ and the algebra $\mathcal{Ol}$ over Ω .

(b) For the purposes of this section we shall define:

$$A \approx B \iff (\neg A \succ B \ \& \ \neg B \succ A) \text{ ;}$$

$$A \sim B \iff \forall C \, (A \approx C \iff B \approx C) \text{ , where}$$

$$A, B, C \in \mathcal{Ol} \text{ .}$$

The relation $\approx$ (called the <u>indifference</u> <u>relation</u>) is reflexive and symmetric, but not transitive. The relation $\sim$ (called the indistinguishability relation) is reflexive, symmetric, transitive, and monotonic; that is $(C \sim A \;\&\; A \succ B) \Longrightarrow C \succ B$.

Sometimes we shall need the set $N(A)$, called the neighborhood of the event A, which is simply the set $\{B \in \mathcal{E} : B \approx A\}$. Note that for $A, B \in \mathcal{E}$, $N(A) = N(B) \Longleftrightarrow A \sim B$. In $\mathcal{E}$ we get an induced weak ordering $\dot{\succ}$:

$$A \dot{\succ} B \Longleftrightarrow A \succ B \qquad \exists_C [B, C \in N(A) \;\&\; C \succ B] \vee$$

$$\exists_D [A, D \in N(B) \;\&\; A \succ D] \;.$$

We shall seldom use these last two notions, even though they are very important in semiordered structures.

In the sequel we shall discuss also the quotient structure $< \Omega/\sim, \mathcal{E}/\sim, \succ/\sim >$ abbreviated by $< \dot{\Omega}, \dot{\mathcal{E}}, \dot{\succ} >$; in this structure $\approx/\sim$ will be written as $\dot{\approx}$.

(c) There is no doubt that the axioms $S_0 - S_4$ are consistent and independent. It is enough to put $\Omega = \{0, 1\}$, $\mathcal{E} = \{A : A \subseteq \Omega\}$, and define $\succ$ in an obvious way. Then this triple becomes a model for the axioms $S_0 - S_4$.

(d) The crucial axioms are S_2 and S_4. Axioms S_1 and S_3 will later impose the so-called normalization condition on the representing measure. S_4 in fact, will be used over and over again; and we need S_1 to prevent the axioms from being satisfied by a trivial structure.

(e) The definition of <u>infinitely</u> <u>additive</u> <u>semiordered</u> <u>qualitative</u> <u>probability</u> <u>sturctures</u>, which can be represented by probability measures on $\mathfrak{A}$ and by jnd-measures (see Theorem 6), does not cause any fundamental difficulties. The axioms (particularly the analogue of axiom S_4) are, however, extremely complicated, and much less intuitive than those given above; this can be checked by a glance at Theorem 3 and Corollary 3. The infinite case will therefore be omitted here. As usual, in this case the topological properties of $\succ$ may be of considerable help in simplifying the solution.

In the following theorem we examine the content of the above definition.

THEOREM 4 <u>Let</u> $< \Omega, \mathfrak{A}, \succ >$ <u>be</u> <u>a</u> FASQP-structure. <u>Then</u> <u>for</u> <u>all</u> $A, B, C, D \in \mathfrak{A}$ <u>the</u> <u>following</u> <u>formulas</u> <u>are</u> <u>satisfied</u>:

(1) $A \succ B \;\&\; C \succ D \implies (A \succ D \vee C \succ B)$;

(2) $A \succ B \;\&\; C \succ A \implies (D \succ B \vee C \succ D)$;

(3) $A \succ B \;\&\; B \succ C \implies (A \succ D \vee D \succ C)$;

(4) $A \succ B \iff A \cup D \succ B \cup D$, if $A, B \perp D$;

(5) $A \succ B \iff \bar{B} \succ \bar{A}$;

(6) $A \subseteq B \implies \neg A \succ B$;

(7) $\neg \emptyset \succ A \;\&\; \neg A \succ \Omega$;

(8) $A \succ B \;\&\; B \succ C \implies A \succ C$;

(9) $A \succ \emptyset \iff \Omega \succ \bar{A}$;

(10) $\neg A \approx \emptyset \Longrightarrow A \succ \emptyset$;

(11) $\neg A \approx \Omega \Longrightarrow \Omega \succ A$;

(12) $\forall_{i < n+1} [A_i \succ B_i] \Longrightarrow B_{n+1} \succ A_{n+1}$, <u>if</u> $\sum_{i \leq n+1} \hat{A}_i = \sum_{i \leq n+1} \hat{B}_i$

<u>and</u> $A_i, B_i \in \mathfrak{A}$, $1 \leq i \leq n+1$;

(13) $A \succ B \;\&\; C \succ D \Longrightarrow A \cup C \succ B \cup D$, if $A \lfloor C \;\&\; B \lfloor D$;

(14) $A \succ B \;\&\; C \succ D \Longrightarrow A \cup C \succ B \cup D$, if $A \lfloor C$;

(15) $A \succ B \Longrightarrow \neg B \succ A$;

(16) $A \cup B \succ C \cup D \Longrightarrow (A \succ C \quad B \succ D)$, if $C \lfloor D \;\&\; AB \succ \emptyset$;

(17) $A \approx \emptyset \;\&\; B \approx \emptyset \Longrightarrow A \approx B$;

(18) $A \subseteq B \;\&\; A \succ \emptyset \Longrightarrow B \succ \emptyset$;

(19) $A \subseteq B \;\&\; B \approx \emptyset \Longrightarrow A \approx \emptyset$;

(20) $A \approx B \Longleftrightarrow \bar{A} \approx \bar{B}$;

(21) $A \approx \Omega \;\&\; B \approx \Omega \Longrightarrow A \approx B$;

(22) $A \succ B \Longleftrightarrow A - B \succ \emptyset$, if $B \subseteq A$;

(23) $A \approx B \Longleftrightarrow A \cup C \approx B \cup C$, if $A, B \lfloor C$;

(24) $A \sim B \Longleftrightarrow \bar{A} \sim \bar{B}$;

(25) $A \subseteq B \;\&\; A \succ C \Longrightarrow B \succ C$;

(26) $A \subseteq B \;\&\; B \sim \emptyset \Longrightarrow A \sim \emptyset$;

(27) $A \subseteq B \;\&\; A \sim \Omega \Longrightarrow B \sim \Omega$;

(28) $A \sim B \Longleftrightarrow A \cup C \sim B \cup C$, if $A, B \lfloor C$;

(29) $A \sim B \;\&\; C \sim D \Longrightarrow A \cup C \sim B \cup D$, if $A \lfloor C \;\&\; B \lfloor D$;

(30) $\langle \dot{\Omega}, \dot{\mathfrak{A}}, \dot{\succ} \rangle$ is FASQP-structure ;

(31) $A \succ B \vee B \succ A \vee A \approx B$, *and each of the formulas excludes the other two*;

(32) $(A \succ B \ \& \ B \approx C \ \& \ C \succ D) \Longrightarrow A \succ D$;

(33) $(A \succ B \ \& \ B \succ C \ \& \ B \approx D) \Longrightarrow (\neg A \approx D \vee \neg C \approx D)$;

(34) $A \dot{\succ} B \Longrightarrow \neg B \dot{\succ} A$;

(35) $A \sim B \Longrightarrow \neg A \dot{\succ} B \ \& \neg B \dot{\succ} A$;

(36) $A \dot{\succ} B \ \& \ B \dot{\succ} C \Longrightarrow A \dot{\succ} C$;

(37) $\langle \dot{\mathfrak{A}}, \dot{\succ} \rangle$ *is a weak ordering structure*.

Proof:

(1) (a) Suppose that $A \succ B \ \& \ C \succ D \ \& \neg A \succ D$. In S_4 put $n = 2$ and $A_1 = A$, $B_1 = B$, $A_2 = C$, $B_2 = D$, $C_1 = A$, $D_1 = D$, $C_2 = C$, $D_2 = B$. Then since $\hat{A}_1 + \hat{A}_2 + \hat{D}_1 + \hat{D}_2 = \hat{B}_1 + \hat{B}_2 + \hat{C}_1 + \hat{C}_2$, we have $C \succ B$.

(b) Suppose that $A \succ B \ \& \ C \succ D \ \& \neg C \succ B$. As before, put $A_1 = A$, $B_1 = B$, $A_2 = C$, $B_2 = D$, $C_1 = C$, $D_1 = B$, $C_2 = A$, $D_2 = D$. Then obviously we get again

$$\hat{A}_1 + \hat{A}_2 + \hat{D}_1 + \hat{D}_2 = \hat{B}_1 + \hat{B}_2 + \hat{C}_1 + \hat{C}_2 .$$

Thus $A \succ D$.

(2) (a) Assume that $A \succ B \ \& \ C \succ A \ \& \neg D \succ B$; put $A_1 = A$, $B_1 = B$, $A_2 = C$, $B_2 = A$, $C_1 = D$, $D_1 = B$, $C_2 = C$, $D_2 = D$. Again the condition on characteristic functions is satisfied. Hence using S_4 we get the conclusion.

(b) Proof is the same as in (a).

(3) Use the same technique as in (2).

(4) Put $A_1 = A$, $D_1 = B \cup D$, $B_1 = B$, $C_1 = A \cup D$. Obviously,

$\hat{A}_1 + \hat{D}_1 = \hat{B}_1 + \hat{C}_1$, since $A, B \perp D$.

Using S_4 we get the conclusion in both directions.

(5) Use S_4 with $n = 1$.

(6) $A \subseteq B$ implies $B = A \cup \bar{A}B$. Now

$\emptyset \cup A \succ A \cup \bar{A}B = B \Longleftrightarrow \emptyset \succ \bar{A}B$ holds in view of (4).

Finally, since $\neg\, \emptyset \succ A$ (as we can check from S_3), we get the conclusion.

(7) If $\emptyset \succ A$, then by S_3, $\emptyset \succ \emptyset$, which is a contradiction.
For the second part use (5), and then the first part of the theorem (7).

(8) Follows from (3) by putting $D = A$.

(9) Use (5).

(10) The assumption implies that $A \succ \emptyset \vee \emptyset \succ A$. In view of (7) we get $A \succ \emptyset$.

(11) Use (7) and the definition of $\approx$.

(12) In S_4 put $C_i = D_i = \emptyset$ for $1 \leq i \leq n-1$, and $D_n = A_{n+1}$,

$C_n = B_{n+1}$. Clearly from the assumption we get

$$\sum_{i \leq n} (\hat{A}_i + \hat{D}_i) = \sum_{i \leq n} (\hat{B}_i + \hat{C}_i) .$$

Naturally we have also $\forall_{i < n}(A_i \succ B_i \;\&\; \neg\, C_i \succ D_i)$ and $A_n \succ B_n$.

Thus, by S_4 we have $B_{n+1} \succ A_{n+1}$.

(13) Follows from (12).

(14) Follows from S_4 ; for

$$(BD)^{\wedge} + \hat{A} + \hat{C} + (B \cup D)^{\wedge} = \hat{B} + \hat{D} + (A \cup C)^{\wedge} + \hat{\emptyset}, \quad \text{if } A \perp C .$$

Now $A \succ B$, $C \succ D$, $\neg\, \emptyset \succ BD$ by the assumption. Hence $A \cup C \succ B \cup D$.

(15) $A \succ B \iff \bar{B} \succ \bar{A}$ by (15). Now if $B \succ A$ were the case, then by (13) we would have $\Omega \succ \Omega$, which is contrary to S_2 . Consequently, $\neg\, B \succ A$.

(16) (a) Clearly $(AB)^{\wedge} + (A \cup B)^{\wedge} + \hat{C} + \hat{D} = (C \cup D)^{\wedge} + \hat{A} + \hat{B} + \hat{\emptyset}$,

if $C \perp D$.

Assume $A \cup B \succ C \cup D$ and $AB \succ \emptyset$, and let $\neg\, A \succ C$. Then by S_4 we get immediately $B \succ D$.

(b) Proof is similar to the proof of (a).

(17) Let $A \approx \emptyset$ & $B \approx \emptyset$ and $\neg\, A \approx B$. Then $A \succ B \vee B \succ A$; so, by S_3 , $A \succ \emptyset$ or $B \succ \emptyset$, which is a contradiction.

(18) $B = A \cup \bar{A}B$. Thus $A \succ \emptyset \iff A \cup \bar{A}B = B \succ \bar{A}B$ by (4). Finally, in view of S_3 , we get $B \succ \emptyset$.

(19) Let $A \subseteq B$ & $B \approx \emptyset$ and $\neg\, A \approx \emptyset$. Then by (10) $A \succ \emptyset$ and by (18) $B \succ \emptyset$, which is a contradiction.

(20) Use the definition of $\approx$, and (5).

(21) Use (5) and (20).

(22) $\hat{B} + (\bar{B}A)^{\wedge} = \hat{A} + \hat{\emptyset}$, if $B \subseteq A$. Use (12) twice.

(23) Use (5) twice.

(24) Use the definition of $\sim$, and (20) .

(25) $A \subseteq B \Longrightarrow \bar{B} \subseteq \bar{A}$, so $\bar{C} \succ \bar{A} \Longrightarrow \bar{C} \succ \bar{B}$ by S_3 . Thus $B \succ C$.

(26) Assume that $A \subseteq B$ & $B \sim \emptyset$, and $\neg A \sim \emptyset$. Then $A \approx B$ & $A \approx \emptyset$ in view of (19) and (17). Since $\neg A \sim B$, we have two cases:

(a) $\exists_C [C \approx A \;\&\; C \succ B]$.

From $A \subseteq B$ it follows by S_3 that $C \succ A$, which is impossible. The case $C \approx A$ & $B \succ C$ would lead to $B \succ \emptyset$.

(b) $\exists_C [C \approx B \;\&\; C \succ A]$.

Hence, $C \succ \emptyset$ and also $C \approx \emptyset$, since $B \sim \emptyset$. But this is a contradiction. The case $C \approx B$ & $A \succ C$ leads to $A \succ \emptyset$ which is also impossible.

(27) $A \sim \Omega \iff \bar{A} \sim \emptyset$ by (24). Use (26) and again (24).

(28) Let $A, B \perp C$. Then

$$A \sim B \iff A/\!\sim \; = B/\!\sim \; \iff A/\!\sim \cup \; C/\!\sim \; = B/\!\sim \cup \; C/\!\sim \; \iff$$

$$A \cup C/\!\sim \; = B \cup C/\!\sim \; \iff A \cup C \sim B \cup C .$$

(29) $A \sim B \implies A/\!\sim \; = B/\!\sim$, $C \sim D \implies C/\!\sim \; = D/\!\sim$. Assuming $A \perp C$ & $B \perp D$, we get $A/\!\sim \cup \; C/\!\sim \; = B/\!\sim \cup \; D/\!\sim$. Hence we have also $A \cup C \sim B \cup D$.

(30) Use the fact that $\sim$ is a congruence relation.

(31) - (37) are trivial consequences of the previous cases. Q. E. D.

Theorem 4 illuminates the intuitive content and the adequacy of our definition. Before we proceed to the formal justification of the definition by proving the so-called <u>Representation Theorem</u>, we shall quote an easy consequence of Theorem 2, due to Scott [11]:

LEMMA 1 *Let $\mathcal{V}$ be a finite-dimensional real linear vector space and let $0 \neq M \subseteq N \subseteq \mathcal{V}$, where N is finite and all its elements have rational coordinates with respect to a given basis; further,* let $N = \{-v : v \in N\}$ (i.e. N is symmetric).

Then there exists a linear functional $\varphi : \mathcal{V} \longrightarrow Re$ such that

$$\varphi(v) \geq 0 \iff v \in M \quad \text{for all} \quad v \in N$$

if and only if

(α) $v \in M$ or $-v \in M$;

(β) $\sum_{i \leq m} v_i = 0 \;\&\; \forall_{i \leq m}(v_i \in M) \quad -v_m \in M$; where

$v, v_i \in N$, $1 \leq i \leq m$.

The proof is given in Scott [11] and for brevity will be omitted here.

THEOREM 5 (Representation Theorem) Let $< \Omega , \mathcal{A} , \succ >$ be a structure, where Ω is a nonempty finite set, $\mathcal{A}$ is the Boolean algebra of subsets of Ω , and $\succ$ is a binary relation on $\mathcal{A}$. Then $< \Omega , \mathcal{A} , \succ >$ is a FASQP-structure if and only if there exists a finitely additive probability measure P and a real number $\mathcal{E}$ such that $< \Omega , \mathcal{A} , P >$ is a probability space and for all $A , B \in \mathcal{A}$:

$$A \succ B \iff P(A) \geq P(B) + \mathcal{E} \quad , \text{ where } 0 < \mathcal{E} \leq 1 ;$$

$$A \sim B \Rightarrow P(A) = P(B) .$$

The theorem remains valid if the representation is given in the form

$$A \succ B \iff P(A) > P(B) + \mathcal{E} \quad , \text{ where } 0 \leq \mathcal{E} < 1 ;$$

$$A \sim B \Rightarrow P(A) = P(B) .$$

If $\mathcal{E} = 0$, then the FASQP-structure reduces to a FAQP-structure (finitely additive qualitative probability structure), such as discussed in Section 2.2.

Proof:

I. The existence of a probability measure P on $\mathcal{A}$ and a real number $\mathcal{E}$.

Suppose that $< \Omega , \mathcal{A}, \succ >$ is a FASQP-structure. Then in view of Theorem 4(30) $< \dot{\Omega} , \dot{\mathcal{A}} , \dot{\succ} >$ is also a FASQP-structure. Let us define $\dot{\Omega}_o = \dot{\Omega} \cup \{e_{m+1}\}$, where $e_{m+1} \notin \dot{\Omega}$, $|\dot{\Omega}| = m$,*) $\dot{\Omega} = \{e_1, e_2, \ldots, e_m\}$, $E = \{e_{m+1}\}$. Then any element A of $\dot{\mathcal{A}}$ **) can be uniquely represented by its characteristic function $\hat{A}$ which we shall consider now as a vector

$$\hat{A} = < \hat{A}(e_1), \hat{A}(e_2), \ldots, \hat{A}(e_m), \hat{A}(e_{m+1}) >$$

in the m+1-dimensional real linear vector space $\mathcal{V}(\dot{\Omega}_o)$, generated by vectors $(\{e_i\})^{\wedge}$, $1 \leq i \leq m+1$. It should be clear what is meant by $\hat{A} + \hat{B}$ and $\alpha \cdot \hat{A}$ in $\mathcal{V}(\dot{\Omega}_o)$, if α is a real number and $A, B \in \dot{\mathcal{A}}$. (For the time being we use the same variables as we used for the elements of $\mathcal{A}$; this is for simplicity of notation.) The reader should consult Remarks (1) given after Theorem 3 in Section 2.2.

*) $|A|$ denotes the cardinality of the set A.

**) Variables A, B, C, D, ... are now running over the algebra $\dot{\mathcal{A}}$.

Let us put

$$N = \{\hat{A} - \hat{B} - \hat{E} : A, B \in \mathfrak{C}\} \cup \{\hat{B} - \hat{A} + \hat{E} : A, B \in \mathfrak{C}\} \text{ and}$$

$$M = \{\hat{A} - \hat{B} - \hat{E} : A, B \in \mathfrak{C} \,\&\, A \succeq B\} \cup \{\hat{B} - \hat{A} + \hat{E} : A, B \in \mathfrak{C} \,\&\, \neg A \succeq B\} .$$

Then surely $\emptyset \neq M \subseteq N \subseteq \mathcal{V}(\dot{\Omega}_0)$; N is finite and symmetric, and contains only rational vectors with respect to the basis $\{((\{e_i\})^{\wedge}\}_{i=1}^{m+1}$. For, $e_{m+1} \in M$ by S_2, $\mathfrak{C}$ is finite, and furthermore, $v \in N \Longleftrightarrow -v \in N$ for any $v \in \mathcal{V}(\dot{\Omega}_0)$.

If $v \in N$, then $v \in M$ or $-v \in M$, since $A \succeq B$ or $\neg A \succeq B$, where $A, B \in \mathfrak{C}$ and $v = \hat{A} - \hat{B} - \hat{E}$ or $v = \hat{B} - \hat{A} + \hat{E}$. Therefore the condition (α) in Lemma 1 is satisfied.

Now the condition $\forall_{i < p}[v_i \in M]$ in (β), Lemma 1, is equivalent to the condition

$$v_i = \hat{A}_i - \hat{B}_i - \hat{E} \,\&\, A_i \succeq B_i \tag{2.1}$$

or

$$v_i = \hat{B}_i - \hat{A}_i + \hat{E} \,\&\, \neg A_i \succeq B_i \; ; \tag{2.2}$$

that is to say, some of the v_i's have the form (2.1) and the rest have the form (2.2). If we relabel the sequence $\{v_i\}_{i=1}^{p-1}$ so that the first k elements $(k < p)$ have the form (2.1) and the remainder the form (2.2), then we get an alternative version of (2.1) and (2.2):

$$\begin{aligned} & v_i = \hat{A}_i - \hat{B}_i - \hat{E} \,\&\, A_i \succeq B_i \, , \quad 1 \leq i \leq k \\ \text{or} \quad & v_i = \hat{B}_i - \hat{A}_i + \hat{E} \,\&\, \neg A_i \succeq B_i \, , \quad k+1 \leq i \leq p-1 \, , \end{aligned} \tag{2.3}$$

where k is some natural number $0 \leq k \leq p-1$.

The condition $\sum_{i \leq p} v_i = 0$ is equivalent to

$$\sum_{i=1}^{k} (\hat{A}_i - \hat{B}_i - \hat{E}) + \sum_{i=k+1}^{p-1} (\hat{B}_i - \hat{A}_i + \hat{E}) + v_p = 0 ,$$

that is,

$$v_p + (p -2k + 1)\cdot\hat{E} + \sum_{1 \leq i \leq k} (\hat{A}_i - \hat{B}_i) +$$

$$\sum_{k+1 \leq i \leq p-1} (\hat{B}_i - \hat{A}_i) = 0 . \tag{2.4}$$

Since $v_p \in N$ we get two cases:

A) $v_p = \hat{A}_p - \hat{B}_p - \hat{E}$:

Since $\hat{E}$ cannot be written as a linear combination of the elements of $\mathcal{V}(\dot{\Omega})$ (which are generated by vectors $\{(\{e_i\})^{\wedge}\}_{i=1}^{m}$ actually belonging to $\{\hat{A} : A \in \dot{\mathcal{U}}\}$), $\hat{E}$ cannot occur in (2.4) the same number of times negated and unnegated. Therefore $(p - 2k - 1)\cdot\hat{E} - \hat{E} = \hat{\emptyset}$, that is $p = 2k + 2$. Thus the equation (2.4) can be rewritten as follows:

$$\hat{A}_p + \sum_{i=1}^{k} \hat{A}_i + \sum_{i=k+1}^{p-1} \hat{B}_i = \hat{B}_p + \sum_{i=1}^{k} \hat{B}_i + \sum_{i=k+1}^{p-1} \hat{A}_i . \tag{2.5}$$

Using the substitution

$A_i^* = A_i$, $B_i^* = B_i$ for $1 \leq i \leq k$,

$$A^*_{k+1} = A_p \ , \quad B^*_{k+1} = B_p \ ,$$

$C_i = B_{i+k}$, $D_i = A_{i+k}$ for $k + 1 = p - k - 1 \geq i \geq 1$,

we got from (2.5) the following formula:

$$\sum_{i=1}^{k+1} (\hat{C}_i + \hat{A}^*_i) = \sum_{i=1}^{k+1} (D_i + \hat{B}^*_i) \ . \tag{2.6}$$

The conditions

$$A_i \dot{\succ} B_i \ (1 \leq i \leq k) \quad \text{and} \quad \neg A_i \dot{\succ} B_i \quad (k+1 \leq i \leq p-1)$$

in (2.3) are now equivalent to the following condition:

$$\forall_{i < k+1} [A^*_i \dot{\succ} B^*_i \ \& \ \neg D_i \dot{\succ} C_i] \ \& \ D_{k+1} \dot{\succ} C_{k+1} \ .$$

Finally, Lemma 1 gives us $-v_p \in M$, that is, $\hat{B}_p - \hat{A}_p + \hat{E} \in M$,

which is equivalent to $A^*_{k+1} \dot{\succ} B^*_{k+1}$.

B) $v_p = \hat{B}_p - \hat{A}_p + \hat{E}$:

For similar reasons as before, $p = 2k$ and then (2.4) becomes

$$\sum_{i=1}^{k} \hat{A}_i + \sum_{i=k+1}^{p} B_i = \sum_{i=1}^{k} \hat{B}_i + \sum_{i=k+1}^{p} \hat{A}_i \ . \tag{2.7}$$

Now if we put

$A^*_i = A_i$, $B^*_i = B_i$ for $1 \leq i \leq k$ and

$C_i = B_{i+k}$, $D_i = A_{i+k}$ for $1 \leq i \leq k = p - k$, then (2.7) becomes

$$\sum_{i=1}^{k} (\hat{A}^*_i + \hat{C}_i) = \sum_{i=1}^{k} (\hat{B}^*_i + \hat{D}_i) \ . \tag{2.8}$$

The conditions $A_i \succeq B_i$ $(1 \leq i \leq k)$ and $\neg A_i \succeq B_i$ $(k+1 \leq i \leq p-1)$ in (2.3) are equivalent to the condition

$$\forall_{i < k} [A_i^* \succeq B_i^* \ \& \ \neg D_i \succeq C_i] \ \& \ A_k^* \succeq B_k^* .$$

Lemma 1 gives us $-v_p \in M$, that is, $\hat{A}_p - \hat{B}_p - \hat{E} \in M$, which is equivalent to $D_k \succeq C_k$.

Finally, joining the cases A) and B) and changing the notation, we get for $p = 2k + 2$, $\sum_{i=1}^{k+1} (\hat{A}_i + \hat{C}_i) = \sum_{i=1}^{k+1} (\hat{B}_i + \hat{D}_i)$; moreover,

$$\forall_{i < k+1} [A_i \succeq B_i \ \& \ \neg D_i \succeq C_i] \text{ implies } [D_{k+1} \succeq C_{k+1} \Longrightarrow A_{k+1} \succeq B_{k+1}] .$$

Similarly, for $p = 2k$ we obtain

$$\sum_{i=1}^{k} (\hat{A}_i + \hat{C}_i) = \sum_{i=1}^{k} (\hat{B}_i + \hat{D}_i) ;$$

and $\forall_{i < k} [A_i \succeq B_i \ \& \ \neg D_i \succeq C_i]$ implies $[A_k \succeq B_k \Longrightarrow D_k \succeq C_k]$,

that is, for $n = p/2$

$$\forall_{i < n} [A_i \succeq B_i \ \& \ C_i \succeq D_i] \ \& \ A_n \succeq B_n \Longrightarrow C_n \succeq D_n , \quad \text{if}$$

$$\sum_{i \leq n} (\hat{A}_i + \hat{D}_i) = \sum_{i \leq n} (\hat{B}_i + \hat{C}_i) .$$

The above reduction of axioms S_2 and S_4 to the conditions (α) and (β) in Lemma 1 allows us to use the conclusion of Lemma 1. That is to say, S_2 and S_4 are the necessary and sufficient conditions for the existence of a linear functional $\varphi : \mathcal{V}(\dot{\Omega}_0) \longrightarrow \mathrm{Re}$ such that $\varphi(v) \geq 0 \Longleftrightarrow v \in M$ for all $v \in N$.

Since $\hat{E} \in M$ (axiom S_2), we have $\varphi(\hat{E}) \geq 0$, and since $-\hat{E} \notin M$, it follows that $\varphi(-\hat{E}) = -\varphi(\hat{E}) < 0$; hence $\varphi(\hat{E}) > 0$.

If $A, B \in \mathcal{O}\mathcal{L}$, then

$A \succ B \Longleftrightarrow \hat{A} - \hat{B} - \hat{E} \in M \Longleftrightarrow \varphi(\hat{A}) \geq \varphi(\hat{B}) + \varphi(\hat{E})$.

Consequently, S_1 gives us $\varphi(\hat{\Omega}) \geq \varphi(\hat{\emptyset}) + \varphi(\hat{E}) > 0$, so that we can put

$$\varphi_0(\hat{A}) = \frac{\varphi(\hat{A})}{\varphi(\hat{\Omega})} .$$

In order to simplify the notation, we translate the result from the vector space $\mathcal{V}^{\varphi}(\dot{\Omega}_0)$ into the Boolean algebra $\mathcal{O}\mathcal{L}$ (c.f. Section 1.4 and Remark (1), given after Theorem 3, in Section 2.2) by putting $\psi(A) = \varphi_0(\hat{A})$. We also define the jnd-measure $\mathcal{E}$ to be $\psi(E)$.

In view of S_1 we have $0 < \mathcal{E} \leq 1$. Obviously

(i) $\psi(\dot{\Omega}) = 1$,

(ii) $A \perp B \Rightarrow \psi(A \cup B) = \psi(A) + \psi(B)$.

Clearly for $A \perp B$ we have $\psi(A \cup B) = \varphi_0((A \cup B)\hat{\ }) =$ $\varphi_0(\hat{A} + \hat{B}) = \varphi_0(\hat{A}) + \varphi_0(\hat{B}) = \psi(A) + \psi(B)$.

After translating into the new notation we get also

(iii) $A \succ B \Longleftrightarrow \psi(A) \geq \psi(B) + \mathcal{E}$.

Now we shall prove that $\psi(A) \geq 0$ for $A \in \dot{\mathfrak{A}}$. Assume that $\psi(A) < 0$ for some $A \in \dot{\mathfrak{A}}$. Obviously $A \stackrel{\cdot}{\approx} \emptyset$ ($A \stackrel{\cdot}{\succ} \emptyset$ would give $\psi(A) \geq \mathcal{E}$, and $\emptyset \stackrel{\cdot}{\succ} A$ is impossible in view of Theorem 4(7)), and $A \neq \emptyset$. Therefore we get two cases:

a) $B \stackrel{\cdot}{\approx} A \;\&\; B \stackrel{\cdot}{\succ} \emptyset$ for some $B \in \dot{\mathfrak{A}}$. Hence $\psi(B) \geq \mathcal{E}$, so that $\psi(B) - \psi(A) > \mathcal{E}$ which means $B \stackrel{\cdot}{\succ} A$. But this is a contradiction. Case $B \stackrel{\cdot}{\approx} A \;\&\; \emptyset \stackrel{\cdot}{\succ} B$ contradicts Theorem 4(7)).

b) $B \stackrel{\cdot}{\approx} \emptyset \;\&\; A \stackrel{\cdot}{\succ} B$ for some $B \in \dot{\mathfrak{A}}$. Thus, in view of S_3 we have $A \stackrel{\cdot}{\succ} \emptyset$, which is impossible. The case $B \stackrel{\cdot}{\approx} \emptyset \;\&\; B \stackrel{\cdot}{\succ} A$ would contradict the consequent of S_3. Hence the assumption $\psi(A) < 0$ leads in all cases to a contradiction. Consequently, we have for $A \in \dot{\mathfrak{A}}$:

(iv) $\psi(A) \geq 0$.

Finally, if we put $P(A) = \psi(A/\sim)$, <u>where now</u> $A \in \mathfrak{A}$, *) then P is a real valued function on $\mathfrak{A}$ and the conditions (i) - (iv) are satisfied if we replace ψ by P and the algebra $\dot{\mathfrak{A}}$ by $\mathfrak{A}$. Moreover,

(v) $A \sim B \Longrightarrow P(A) = P(B)$, if $A, B \in \mathfrak{A}$.

Thus on the basis of (i) - (v), $< \Omega, \mathfrak{A}, P >$ is a probability space, and P is the desired finitely additive probability measure of Theorem 5.

II. The probability measure P on $\mathfrak{A}$ and the existence of a real number $\mathcal{E}$ $(0 < \mathcal{E} \leq 1)$ imply the axioms $S_1 - S_4$.

---

*) Variables A, B, C, D, ... are now running over $\mathfrak{A}$ again.

Let $< \Omega , \mathfrak{A} , P >$ be a probability space such that

$A \succ B \Leftrightarrow P(A) \geq P(B) + \varepsilon$, where $0 < \varepsilon \leq 1$, and

$A \sim B \Rightarrow P(A) = P(B)$, for all $A, B \in \mathfrak{A}$.

One can easily check that:

$1 = P(\Omega) \geq \varepsilon$ implies S_1 ;

$\neg P(A) \geq P(A) + \varepsilon$ implies S_2 ;

$A \subseteq B \Rightarrow P(A) \leq P(B)$ and $P(C) \geq P(B) + \varepsilon \geq P(A) + \varepsilon$

together imply S_3 ; and finally, if we put $\varphi_0((\{e_i\})^\wedge)$ for

$1 \leq i \leq m$, we get the linear functional from Lemma 1.

The condition

$$\sum_{i \leq n} (\hat{A} + \hat{D}) = \sum_{i \leq n} (\hat{B} + \hat{C})$$

then implies

$$\sum_{i \leq n} [\psi(A_i) + \psi(D_i)] = \sum_{i \leq n} [\psi(B_i) + \psi(C_i)] \tag{2.9}$$

and thus the condition

$$\forall_{i < n} [A_i \succ B_i \ \& \ \neg C_i \succ D_i] \ \& \ A_n \succ B_n$$

gives us

$$\sum_{i < n} P(A_i) \geq \sum_{i < n} P(B_i) + (n - 1) \cdot \varepsilon \quad ;$$

$$\sum_{i < n} P(C_i) \geq \sum_{i < n} P(D_i) + (n - 1) \cdot \mathcal{E} \quad ;$$

and

$$P(A_n) \geq P(B_n) + \mathcal{E} \quad .$$

Adding together these inequalities and subtracting equality (2.9) from the result, we get

$$P(C_n) > P(D_n) + \mathcal{E} \quad ,$$

which implies $C_n \succ D_n$.

Thus the proof of the Representation Theorem is complete.

A question arises of what group or set T of transformations the probability measure in Theorem 5 is unique up to; or in other words, 'how many' different probability measures can we have, once a binary relation $\succ$ in a FASQP-structure is given. We might expect some periodic functions with period $\mathcal{E}$ to be the elements of the unknown set T . The complete answer does not seem to be simple, and we therefore leave it as an open problem. Some further discussion of this subject will be given in Chapter 5.

We pointed out that the intransitivity of the relation $\approx$ reflects the inability of a measurement method (or apparatus) to distinguish or recognize two different magnitudes of the measured quantity, when their difference is below the sensitivity of the method. More refined measurement methods are needed to make these magnitudes distinguishable. This amounts to considering a lattice of relations $\{\approx_i\}_{i\in I}$ where $\approx_i$ is <u>finer</u> than $\approx_j$ iff $A \approx_i B \Longrightarrow A \approx_j B$ for all $A, B \in \mathcal{U}$ $(i, j \in I)$. The set $\{\approx_i\}_{i\in I}$ then characterizes the class of measurement methods from the point of view of sensitivity.

In psychology there are problems in which $\approx$ is not constant, but varies with the entity on which the measurement is performed. In particular, if $A \in \mathcal{U}$, then $\underline{\mathcal{E}}(A)$ and $\overline{\mathcal{E}}(A)$ characterize the <u>change</u> in probability necessary for indifference $\approx$ to become preference $\succ$.

If we put for $A \in \mathcal{U}$

$$\begin{aligned} \overline{\mathcal{E}}(A) &= \operatorname{Max}\{P(B) - P(A) : A \approx B \ \&\ B \in \mathcal{U}\}; \\ \underline{\mathcal{E}}(A) &= \operatorname{Max}\{P(A) - P(B) : A \approx B \ \&\ B \in \mathcal{U}\}, \end{aligned} \tag{2.10}$$

then

(i) $0 \leq \overline{\mathcal{E}}(A), \ \underline{\mathcal{E}}(A) < 1$;

(ii) $A \approx B \Longrightarrow P(B) - \underline{\mathcal{E}}(B) \leq P(A) \leq P(B) + \overline{\mathcal{E}}(B)$;

(iii) $A \succ B \Longleftrightarrow P(A) > P(B) + \overline{\mathcal{E}}(B)$;

(iv) $P(A) \leq P(B) + \overline{\mathcal{E}}(B) \Longleftrightarrow P(A) \leq P(B) + \underline{\mathcal{E}}(A)$;

(v) $P(A) < P(B) \Longrightarrow [P(A) + \overline{\mathcal{E}}(A) < P(B) + \overline{\mathcal{E}}(B) \vee$
$P(A) + \underline{\mathcal{E}}(B) < P(B) + \underline{\mathcal{E}}(A)]$;

(vi) $A \subseteq B \Longrightarrow \overline{\mathcal{E}}(A) \leq \overline{\mathcal{E}}(B)$.

This is easily checked. Consequently, Theorem 5 can be proven also for the _variable_ _jnd's_ $\overline{\mathcal{E}}(A)$, $\underline{\mathcal{E}}(A)$, given in (2.10). It is an open problem how to give a representation of $< \Omega , \mathcal{O}\!\mathcal{L} , \succ >$ in terms of P, $\underline{\mathcal{E}}$, $\overline{\mathcal{E}}$, without assuming the condition (2.10) _a priori_.

2.4. Quadratic Qualitative Probability Structures

In [34] Luce and Tukey gave a formal presentation of what they called _conjoint measurement structures_. Such structures are linear. Here, by constrast, _nonlinear_ (_quadratic_) measurement structures will be introduced for probability. More concretely, given a finite Boolean algebra $\mathcal{O}\!\mathcal{L}$ of subsets of Ω and a binary relation $\precsim$ *) on the set of Cartesian products of elements from $\mathcal{O}\!\mathcal{L}$, we shall give the necessary and sufficient conditions for the existence of a probability measure P on $\mathcal{O}\!\mathcal{L}$ such that for all $A, B, C, D \in \mathcal{O}\!\mathcal{L}$

$$A \times B \precsim C \times D \Longleftrightarrow P(A) \cdot P(B) \leq P(C) \cdot P(D) .$$

As will be seen later, the appearance of Cartesian products $A \times B$, $C \times D$ here is not essential; we could as well consider the ordered couples $< A, B >$, $< C, D >$. Structures of this

*) For typographical simplicity, we use the same symbol that was used in Section 2.2 for a different ordering.

sort differ from Luce's conjoint measurement structures in three respects: they are finite, the representing function has a special property, namely, it is additive, and finally, the representation is quadratic and not linear. Since most of the laws of classical physics can be represented (using the so-called π-theorem) by equations between a given (additive) empirical quantity and the product of other (additive) empirical quantities (possibly with rational exponents), such a structure is of basic importance in algebraic measurement theory.

For instance, for Ohm's law we might hope to give, for the system of current sources $\{c_i\}_{i \leq n}$ and resistors $\{r_i\}_{i \leq m}$, a representation theorem in the form:

$$< c_i, r_i > \precsim < c_j, r_j > \iff I_i \cdot R_i \leq I_j \cdot R_j ,$$

where on the right we have well-known physical quantities, namely, current and resistance $(i \leq n, \quad j \leq m)$.

This is a digression. Returning to quadratic probability structures, the reader may wonder in what way the formula $A \times B \precsim C \times D \quad (A, B, C, D \in \mathcal{U})$ in (2.12) can be interpreted.

There are several partial interpretations which will be discussed in the sequel:

(a) Qualitative probabilistic independence relation $\perp\!\!\!\perp$:

$$A \perp\!\!\!\perp B \iff AB \times \Omega \sim A \times B ,$$

where, as usual, $A, B \in \mathfrak{A}$ and $\sim$ is the standard equivalence relation induced by $\precsim$.

(b) <u>Qualitative conditional probability relation</u> $\precsim$:

$$A/B \precsim C/D \Longleftrightarrow AB \times D \precsim CD \times B \ , \quad \text{if} \quad \emptyset \times \Omega \prec B \times D \ ,$$

where $A, B, C, D \in \mathfrak{A}$ and $\prec$ is the <u>strict</u> counterpart of $\precsim$. The entities A/B, C/D can be considered here as primitive.

(c) <u>Relevance (positive and negative dependence) relations</u> C_+, C_- :

$$A \; C_+ \; B \Longleftrightarrow A \times B \prec AB \times \Omega \ ;$$

$$A \; C_- \; B \Longleftrightarrow AB \times \Omega \prec A \times B \ ,$$

where $A, B \in \mathfrak{A}$. These notions may be of some help in analyzing causality problems. It is immediately obvious that $A \; C_+ \; B \Leftrightarrow A/\Omega \prec A/B$ and $A \; C_- \; B \Leftrightarrow A/B \prec A/\Omega$.

(d) <u>Qualitative conditional independence relation</u> $\|$:

$$A/C \,\|\, B/C \Longleftrightarrow AC \times BC \sim ABC \times C \ , \quad \text{if} \quad \emptyset \prec C \ ,$$

where $A, B, C \in \mathfrak{A}$.

Since, as can be seen, there are several important interpretations of the formula $A \times B \precsim C \times D$, we shall study the structure of the 'quadratic' relation $\precsim$ in considerable detail.

DEFINITION 2 *A triple* $< \Omega , \mathcal{E}\mathcal{X} , \precsim >$ *is said to be a finitely additive quadratic qualitative probability structure* (FAQQP-structure) *if and only if the following conditions are satisfied:*

Q_0 Ω *is a nonempty finite set;* $\mathcal{E}\mathcal{X}$ *is the Boolean algebra of subsets of* Ω ; *and* $\precsim$ *is a binary relation on* $\{A \times B : A \in \mathcal{E}\mathcal{X} \ \& \ B \in \mathcal{E}\mathcal{X}\}$.

Q_1 $\emptyset \times \Omega \prec \Omega \times \Omega$;

Q_2 $\emptyset \times A \precsim B \times C$;

Q_3 $A \times B \precsim B \times A$;

Q_4 $A \times B \precsim C \times D \vee C \times D \precsim A \times B$;

Q_5 $\forall_{i < n}(A_i \times B_i \precsim A_{\alpha_i} \times B_{\beta_i}) \Longrightarrow A_{\alpha_n} \times B_{\beta_n} \precsim A_n \times B_n$;

Q_6 $\forall_{i < n}(C_i \times D_i \precsim E_i \times F_i) \Longrightarrow E_n \times F_n \precsim C_n \times D_n$;

where $\forall_{i < n}(\emptyset \times \Omega \prec A_i \times B_i)$; $\sum_{i \le n} (C_i \times D_i)^{\wedge} = \sum_{i \le n} (E_i \times F_i)^{\wedge}$;

$A, B, C, D, A_i, B_i, C_i, D_i, E_i, F_i \in \mathcal{E}\mathcal{X}$ $(i \le n)$; α, β *are permutations on* $\{1, 2, \ldots, n\}$, *and* $(C \times D)^{\wedge}$ *denotes the characteristic function of the set* $C \times D$.

Remarks:

(i) We define

$$A \prec B \Longleftrightarrow A \times \Omega \precsim B \times \Omega ;$$

$$A \times B \prec C \times D \Longleftrightarrow \neg\, C \times D \precsim A \times B ;$$

$$A \times B \sim C \times D \Leftrightarrow A \times B \precsim C \times D \ \& \ C \times D \precsim A \times B \ ;$$

$$(A \times B)^{\wedge}(\omega_1, \omega_2) = 1 \ , \ \text{if} \ \omega_1 \in A \ \& \ \omega_2 \in B \ ;$$

$$\text{otherwise} \quad (A \times B)^{\wedge}(\omega_1, \omega_2) = 0 \quad (\omega_1, \omega_2 \in \Omega) \ .$$

(ii) The formula concerning characteristic functions in axiom C_6 can easily be translated into a system of identities among sets; a similar transformation was made in the case of the qualitative probability axioms listed in Section 2.2. Thus the axioms for $\precsim$ contain as primitives <u>only</u> the relation $\precsim$ and the algebra $\mathcal{U}$.

The content of the above definition is laid bare in the following easily proved theorem.

THEOREM 6 <u>Let</u> $< \Omega , \mathcal{U} , \precsim >$ <u>be</u> <u>a</u> FAQQP-structure. <u>Then</u> <u>the</u> <u>following</u> <u>formulas</u> <u>are</u> <u>valid</u> <u>for</u> <u>all</u> A, B, C, D, E, F $\in \mathcal{U}$:

(1) $A \times B \sim A \times B$;

(2) $A \times B \sim B \times A$;

(3) $A \times B \precsim C \times D \ \& \ C \times D \precsim E \times F \Rightarrow A \times B \precsim E \times F$;

(4) $A \times C \precsim B \times C \Leftrightarrow A \times D \precsim B \times D$, if $\emptyset \times \Omega \prec C \times D$;

(5) $A \times B \precsim C \times D \ \& \ E \times C \precsim F \times B \Rightarrow A \times E \precsim F \times D$, if $\emptyset \times \Omega \prec B \times C$;

(6) $A \times B \precsim C \times D \Leftrightarrow B \times A \precsim D \times C$;

(7) $\Omega \times A \precsim B \times \Omega \ \& \ C \times \Omega \precsim \Omega \times D \Rightarrow A \times C \precsim B \times D$;

(8) $A \times \Omega \precsim B \times \Omega \Leftrightarrow A \times A \precsim B \times B$;

(9) $A \times B \sim C \times D \Rightarrow (A \precsim C \Leftrightarrow D \precsim B)$;

(10) $(A \times A \precsim F \times F \ \& \ A \times E \precsim D \times D \ \& \ E \times E \precsim D \times F) \rightarrow A \times E \precsim D \times F$;

(11) $A \precsim B \Leftrightarrow A \times \bar{B} \precsim \bar{A} \times B$;

(12) $\emptyset \prec A \;\&\; \emptyset \prec B \Rightarrow \emptyset \times \Omega \prec A \times B$;

(13) $$\forall_{i \leq n} (A_i \times B_i \precsim C_i \times D_i) \;\&\; \forall_{i \leq n} (C_{\gamma_i} \times D_{\delta_i} \precsim A_{\alpha_i} \times B_{\beta_i})$$

$$\Rightarrow A_{\alpha_n} \times B_{\beta_n} \precsim C_{\gamma_n} \times D_{\delta_n} \;, \quad \text{if} \quad \forall_{i < n} (\emptyset \times \Omega \prec C_{\gamma_i} \times D_{\delta_i}) \;,$$

where $A_i, B_i, C_i, D_i \in \mathcal{Ol}$ $(i \leq n)$, and $\alpha, \beta, \gamma, \delta$ are permutations on $\{1, 2, \ldots, n\}$;

(14) If $A \precsim_0 B \Leftrightarrow A \times \Omega \precsim B \times \Omega$, then $< \Omega , \mathcal{Ol} , \precsim_0 >$ is a FQP-structure.

Theorem 6 will be useful in several ways. In particular, the properties of $\|$ will be derived from it.

Before we proceed to the representation theorem for FAQQP-structures, we must give a brief review of tensor products of ordered vector spaces.

The tensor product of two vector spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ is, roughly speaking, the set of formal sums

$$\sum_{i \leq n} \alpha_i (v_i \otimes w_i) \;, \quad \text{where} \quad \alpha_i \in Re \;, \quad v_i \in \mathcal{V}_1 \;, \quad w_i \in \mathcal{V}_2$$

for $i \leq n$;

this is made into a vector space by considering the following formulas to be valid for all $v, v_1, v_2 \in \mathcal{V}_1$, $w, w_1, w_2 \in \mathcal{V}_2$ and $\alpha \in Re$:

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \;;$$

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2 ;$$

$$\alpha (v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w) .$$

If $\mathcal{V}_1$ and $\mathcal{V}_2$ are equipped with cones $\mathcal{C}_1$, $\mathcal{C}_2$ respectively, inducing orderings in $\mathcal{V}_1$ and $\mathcal{V}_2$, we shall call the couple $< \mathcal{V}_1 \otimes \mathcal{V}_2, \mathcal{C} >$ a <u>tensor product of the ordered vector spaces</u> $< \mathcal{V}_1, \mathcal{C}_1 >$ <u>and</u> $< \mathcal{V}_2, \mathcal{C}_2 >$ iff $\mathcal{C} = \{ \sum_{i \leq n} \alpha_i (v_i \otimes w_i) : v_i \in \mathcal{C}_1 \;\&\; w_i \in \mathcal{C}_2 \;\&\; \alpha_i \in Re$ for $i \leq n \}$.

The tensor product of ordered vector spaces is again an ordered vector space; and, in particular, if $< \mathcal{V}_1, \precsim_1 >$, $< \mathcal{V}_2, \precsim_2 >$ are the given ordered vector spaces and $< \mathcal{V}_1 \otimes \mathcal{V}_2, \precsim >$ is their 'ordered' tensor product, then

(i) $0 \precsim_1 v \;\&\; 0 \precsim_2 w \rightarrow 0 \precsim v \otimes w$;

(ii) $v_1 \otimes w_1 = v_2 \otimes w_2 \rightarrow (v_1 \precsim_1 v_2 \Leftrightarrow w_2 \precsim_2 w_1)$,

where $v, v_1, v_2 \in \mathcal{V}_1$, $w, w_1, w_2 \in \mathcal{V}_2$.

The well-known natural isomorphism between the space of bilinear functionals on $\mathcal{V}_1 \times \mathcal{V}_2$ and the space of linear functionals on $\mathcal{V}_1 \otimes \mathcal{V}_2$, $\mathcal{B}(\mathcal{V}_1, \mathcal{V}_2) \cong (\mathcal{V}_1 \otimes \mathcal{V}_2)$, turns here into an isomorphism between the space of order-preserving bilinear functionals and the space of order-preserving linear functionals.

For finite dimensional ordered vector spaces $< \mathcal{V}_1 , \mathcal{E}_1 >$, $< \mathcal{V}_2 , \mathcal{E}_2 >$, $\dim (\mathcal{V}_1 \otimes \mathcal{V}_2) = \dim \mathcal{V}_1 \cdot \dim \mathcal{V}_2$; if $< \mathcal{V}_1 \otimes \mathcal{V}_2 , \mathcal{E} >$ is the 'ordered' tensor product of $\mathcal{V}_1$ and $\mathcal{V}_2$, then

$$\mathcal{E} = \{ \sum_{i \leq n} v_i \otimes w_i : \forall f \in \mathcal{E}_1^* \; \forall g \in \mathcal{E}_2^* \sum_{i \leq n} f(v_i) \cdot g(w_i) \geq 0 \} ,$$

where $\mathcal{E}_i^*$ denotes the dual cone in $\mathcal{V}_i^*$ and $\mathcal{V}_i^*$ is the dual vector space of $\mathcal{V}_i$ for $i = 1, 2$.

THEOREM 7 (Representation Theorem) Let $< \Omega , \mathcal{O}\!\mathcal{L} , \precsim >$ be a structure, where Ω is a nonempty finite set; $\mathcal{O}\!\mathcal{L}$ is the Boolean algebra of subsets of Ω , and $\precsim$ is a binary relation on $\{A \times B : A \in \mathcal{O}\!\mathcal{L} \;\&\; B \in \mathcal{O}\!\mathcal{L} \}$.

Then $< \Omega , \mathcal{O}\!\mathcal{L} , \precsim >$ is a FAQQP-structure if and only if there exists a finitely additive probability measure P such that $< \Omega , \mathcal{O}\!\mathcal{L} , P >$ is a probability space, and for all $A, B, C, D \in \mathcal{O}\!\mathcal{L}$,

$$A \times B \precsim C \times D \Longleftrightarrow P(A) \cdot P(B) \leq P(C) \cdot P(D) .$$

Proof:

I. Sufficiency

(a) Translation of the problem from the language of relations into geometric language.

We shall first represent the Boolean elements $A \in \mathfrak{A}$ by vectors $\hat{A} = \langle \hat{A}(\omega_1), \hat{A}(\omega_2), \ldots, \hat{A}(\omega_n) \rangle$, where $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $|\Omega| = n$, and $\hat{A}(\omega) = 1$, if $\omega \in A$, $\hat{A}(\omega) = 0$ otherwise. Defining $\hat{A} + \hat{B}$, $\alpha \cdot \hat{A}$ in an obvious way, we generate a vector space $\mathcal{V}(\Omega) = \mathcal{V}$, where $\{\hat{A} : A \in \mathfrak{A}\} \subseteq \mathcal{V}$ and $\dim \mathcal{V} = n$. Defining $\hat{A} \precsim \hat{B} \Longleftrightarrow A \precsim B$, we can generate a cone $\mathcal{C}$ in $\mathcal{V}$ by using the set $\{\hat{B} - \hat{A} : A \precsim B \;\&\; A, B \in \mathfrak{A}\}$; this furnishes $\mathcal{V}$ with an ordering structure, corresponding in a one-one way to the ordering in $\mathfrak{A}$.

The Cartesian product $A \times B$ will be represented by the tensor product $\hat{A} \otimes \hat{B}$ in $\mathcal{V} \otimes \mathcal{V}$.

Putting $\hat{A} \otimes \hat{B} \precsim \hat{C} \otimes \hat{D} \Longleftrightarrow A \times B \precsim C \times D$, we get an ordering on $\mathcal{V} \otimes \mathcal{V}$. This completes the translation.

(b) Translation of the problem from geometric language into functional language.

Translating Q_4 and Q_6 into geometric language of tensors and using Corollary 5, we have the necessary and sufficient conditions for the existence of a linear functional $\psi : \mathcal{V} \otimes \mathcal{V} \longrightarrow \mathrm{Re}$ such that

$$\hat{A} \otimes \hat{B} \precsim \hat{C} \otimes \hat{D} \Longleftrightarrow \psi(\hat{A} \otimes \hat{B}) \leq \psi(\hat{C} \otimes \hat{D}),$$

for all $A, B, C, D \in \mathfrak{A}$.

In view of the isomorphism of the space of positive linear functionals on $\mathcal{V} \otimes \mathcal{V} : \mathcal{L}(\mathcal{V} \otimes \mathcal{V}) \cong \mathcal{B}(\mathcal{V}, \mathcal{V})$, we can pick up a bilinear functional $\varphi : \mathcal{V} \times \mathcal{V} \longrightarrow \mathrm{Re}$, corresponding to ψ and put

$$\hat{A} \otimes \hat{B} \precsim \hat{C} \otimes \hat{D} \Longleftrightarrow \varphi(\hat{A}, \hat{B}) \leq \varphi(\hat{C}, \hat{D})$$

for all $A, B, C, D \in \mathcal{Ct}$.

Now Q_2 compels φ to be non-negative: $\varphi(\hat{A}, \hat{B}) \geq \psi(\hat{\emptyset}, \hat{C}) = 0$ on $\{\hat{A} \otimes \hat{B} : A, B \in \mathcal{Ct}\}$; Q_1 allows us to normalize φ : $\varphi(\hat{\Omega}, \hat{\Omega}) > 0$; and Q_3 forces φ to be symmetric: $\varphi(\hat{A}, \hat{B}) = \varphi(\hat{B}, \hat{A})$.

The last step remains, but it is an important one. It is to show that φ can be split into a product of two linear functionals: $\varphi(\hat{A}, \hat{B}) = f(\hat{A}) \cdot g(\hat{B})$. It is an elementary fact from linear algebra that this can be done if and only if the rank of φ is equal to one. As $\mathcal{V}(\Omega) \otimes \mathcal{V}(\Omega) \cong \mathcal{V}(\Omega \times \Omega)$, this can be expressed also in terms of the matrix of φ . Because of the symmetry of φ , f must be equal to g . Axiom Q_5, translated into geometric language, determines the values of $\varphi(\hat{A}, \hat{B})$ on a system of curves which nowhere intersect each other, as one can check from Theorem 6(4,5), and from countably many similar consequences of Q_5. Since φ is symmetric and linear with respect to each of the arguments, the curves must form a system of symmetric hyperbolas (cf. Aczél, Pickert, and Radó [35]). In fact, since $\psi \in (\mathcal{V} \otimes \mathcal{V})^* \cong \mathcal{V}^* \otimes \mathcal{V}^*$, $\psi = \sum_{i \leq n} f_i \otimes g_i$, where

$f_i, g_i \in \mathcal{V}^*$ for $i \leq n$. Thus $\psi(\hat{A} \otimes \hat{B}) = \sum_{i \leq n} f_i(\hat{A}) \cdot g_i(\hat{B})$.

In view of the above argument, $f_i = \alpha_{i,j} \cdot f_j$, $g_i = \beta_{i,j} \cdot g_j$, where $\alpha_{i,j}$, $\beta_{i,j} \in Re$ for $i,j \leq n$; and the normalization of ψ implies $f = f_i = f_j$; $g = g_i = g_j$ for $i,j \leq n$; and thanks to symmetry, we further get $f = g$ as stated above. Hence we get for all $A, B, C, D \in \mathcal{U}$

$$\hat{A} \otimes \hat{B} \precsim \hat{C} \otimes \hat{D} \Longleftrightarrow f(\hat{A}) \cdot f(\hat{B}) \leq f(\hat{C}) \cdot f(\hat{D}) .$$

(c) Translation of the problem from functional language back to the language of relations.

We switch from $\hat{A} \in \mathcal{V}$ and $f : \mathcal{V} \longrightarrow Re$ to $A \in \mathcal{U}$ and $P : \mathcal{U} \longrightarrow Re$ by translating tensors $\hat{A} \otimes \hat{B}$ into Cartesian products $A \times B$ and putting

$$P(A) = \frac{f(\hat{A})}{f(\hat{\Omega})} \qquad \text{for all } A \in \mathcal{U} \quad .$$

Then clearly $< \Omega , \mathcal{U} , P >$ is a probability space and (2.12) is satisfied.

II. Necessity.

It is a routine matter to show that the axioms $Q_0 - Q_6$ in definition 2 are necessary. Q. E. D.

It should perhaps be pointed out that FAQQP-structures exemplify an important class of fintie quadratic measurement structures not previously discussed in the literature.

2.5. Probabilistically Independent Events

As is well-known, probabilistically independent events play an essential role in the definitions of information and entropy. The independence relation between events is defined entirely in terms of the probability measure P : $P(AB) = P(A) \cdot P(B)$. One wonders whether it is possible to give a definition of a corresponding binary relation $\parallel$ on $\mathcal{E}$ in terms of the qualitative probability relation $\precsim$ on $\mathcal{E}$. It is trivial to see that this is not possible in terms of FQP-structures, but, as has been pointed out, such a relation can be defined in terms of FAQQP-structures by putting

$$A \parallel B \Longleftrightarrow AB \times \Omega \sim A \times B \quad \text{for all} \quad A, B \in \mathcal{E} \quad . \tag{2.13}$$

This definition not only is important for qualitative information and entropy structures, but also can be relevant in applied probability theory, where one does not care too much about the underlying probability structure $< \Omega , \mathcal{E}, P >$, but emphasizes rather the analytic properties of random variables. Under these circumstances the independent random variables could be handled using the basic properties of $\parallel$, without explicit reference to the probability measure P that satisfies the condition

$$A \parallel B \Leftrightarrow P(AB) = P(A) \cdot P(B) .$$

In this section we state a theorem about the basic properties of $\parallel$.

THEOREM 8 *If* $< \Omega , \mathscr{E} , \precsim >$ *is a* FAQQP-structure, *then given* (2.13) *the following formulas are valid when all variables run over* $\mathscr{E}$:

(1) $\emptyset \parallel A$;

(2) $\Omega \parallel A$;

(3) $A \parallel A \Longleftrightarrow (A \sim \Omega \vee A \sim \emptyset)$;

(4) $A \parallel A \Longrightarrow A \parallel B$;

(5) $A \parallel B \,\&\, A \perp B \Longrightarrow (A \sim \emptyset \vee B \sim \emptyset)$;

(6) $A \parallel B \,\&\, A \subseteq B \Longrightarrow (A \sim \emptyset \vee B \sim \Omega)$;

(7) $A \parallel B \,\&\, A \sim B \Longrightarrow \bar{A}B \sim A\bar{B}$;

(8) $A \parallel B \Longleftrightarrow B \parallel A$;

(9) $A \parallel B \Longleftrightarrow A \parallel \bar{B}$;

(10) $A \parallel B \Longleftrightarrow \bar{A} \parallel \bar{B}$;

(11) $A \parallel B \Longrightarrow AB \prec B$, if $A \prec \Omega \,\&\, \emptyset \prec B$;

(12) $A \parallel B \Longrightarrow (\emptyset \prec A \,\&\, \emptyset \prec B \Longrightarrow \emptyset \prec AB)$;

(13) $A \parallel B \,\&\, B \parallel C \Longrightarrow (AB \parallel C \Longleftrightarrow A \parallel BC)$;

(14) $A \parallel B \,\&\, C \parallel D \Longrightarrow (A \precsim C \,\&\, B \precsim D \Longrightarrow AB \precsim CD)$;

(15) $A \parallel B \,\&\, A \parallel C \Longrightarrow A \parallel B \cup C$, if $B \perp C$;

(16) $A \parallel B \,\&\, A \parallel C \Longrightarrow A \parallel B \cap C$, if $B \cup C = \Omega$;

(17) $A \parallel B \,\&\, A \parallel C \Longrightarrow (B \precsim C \Longleftrightarrow A \cup B \precsim A \cup C)$, if $A \prec \Omega$;

(18) $A \parallel B \,\&\, A \parallel C \Longrightarrow (AB \precsim AC \Longleftrightarrow B \precsim C)$, if $\emptyset \prec A$;

(19) $A \sim C \,\&\, AB \sim CB \Longrightarrow (A \parallel B \Longleftrightarrow C \parallel B)$;

(20) $\forall_{i < n}(A_iB_i \precsim A_{\alpha_i}B_{\beta_i}) \Longrightarrow A_{\alpha_n}B_{\beta_n} \precsim A_nB_n$, *if*

$\forall_{i \leq n}(A_i \parallel B_i \,\&\, A_{\alpha_i} \parallel B_{\beta_i} \,\&\, 0 \prec A_iB_i)$, *and* α, β *are permutations on* $\{1, 2, \ldots, n\}$.

The proof is a routine application of Theorem 6. We shall use this theorem throughout Chapter 3. It is rather disappointing that the qualitative independence relation $\perp\!\!\!\perp$, which plays a central role in probability theory, has such complicated properties.

It was Marczewski [36] who argued that probabilistic independence has a different nature from the notions of algebraic, logical, and set-theoretic independence. The fact that this is not precisely true was demonstrated by Maeda [37].

The independence relation $\perp\!\!\!\perp$ can be extended to any (finite) family of events $\{A_i\}_{i \in I} \subseteq \mathcal{U}$ with more than two elements in such a way that the following equivalence is preserved:

$$\{A_i\} \underset{i \in I}{\perp\!\!\!\perp} \Longleftrightarrow \underset{\emptyset \neq I_0 \subseteq I}{\forall} [P(\bigcap_{i \in I_0} A_i) = \prod_{i \in I_0} P(A_i)] .$$

It is sufficient to put

(i) $\{A, B\} \perp\!\!\!\perp \Longleftrightarrow A \perp\!\!\!\perp B$;

(ii) $\{A_i\} \underset{i \in I}{\perp\!\!\!\perp} \Longleftrightarrow \underset{\emptyset \neq I_0 \subset I}{\forall} [\{A_i\} \underset{i \in I_0}{\perp\!\!\!\perp} \ \& \bigcap_{i \in I_0} A_i \perp\!\!\!\perp \bigcap_{i \in I - I_0} A_i]$.

It can be shown easily that

(1) $\{A_i\} \underset{i \in I}{\perp\!\!\!\perp} \Longrightarrow A_i \underset{\substack{i, j \in I \\ i \neq j}}{\perp\!\!\!\perp} A_j$;

(2) $\{A_i\} \underset{i \in I}{\perp\!\!\!\perp} \Longrightarrow \{B_i\} \underset{i \in I}{\perp\!\!\!\perp}$, if $\underset{i \in I}{\forall} [B_i = A_i \vee B_i = \overline{A}_i]$;

(3) $\{A_i\} \underset{i\in I}{\perp\!\!\!\perp} \Longleftrightarrow \underset{J \subseteq I}{\forall} [\{A_i\} \underset{i\in J}{\perp\!\!\!\perp}]$;

(4) $\{A_i\} \underset{i\in I}{\perp\!\!\!\perp} \Longrightarrow \underset{I_1, I_2 \subseteq I \,\&\, I_1 \perp I_2}{\forall} [\bigcap_{i\in I_1} A_i \perp\!\!\!\perp \bigcap_{i\in I_2} A_i]$.

We shall not need these rather general properties; further details about $\underset{i\in I}{\perp\!\!\!\perp}$ are therefore omitted.

Perhaps we should point out that the I-place relation $\underset{i\in I}{\perp\!\!\!\perp}$ enables us to treat probabilistic independence in lattice-theoretic terms. In particular, the lattice-theoretic notion of independence coincides under certain reasonable conditions with the probabilistic relation $\perp\!\!\!\perp$. As mentioned before, this is contrary to what Marczewski [36] maintains.

2.6. Qualitative Conditional Probability Structures

The first part of this section is devoted to the study of some simple algebraic features of the relational structure $< \mathfrak{A}, \precsim >$, where for A, B, C, D $\in \mathfrak{A}$, $A/B \precsim C/D$ is interpreted probabilistically as follows: event A given event B is not more probable than event C given event D. As stated in Section 2.1, the conditional event A/B is defined set-theoretically as an element of the quotient Boolean algebra $\mathfrak{A}/B = \mathfrak{A}/\nabla(B)$, where $\nabla(B)$ is the filter generated by the nonempty event B.

We shall be concerned with problem (P_3) of Section 1.1, the basic interplay between the qualitative conditional probability structure $<\Omega, \mathcal{C}, \precsim>$ and the probability space $<\Omega, \mathcal{C}, P>$. In particular, a representation theorem is proved.

DEFINITION 3 A triple $<\Omega, \mathcal{C}, \precsim>$ is a finite qualitative conditional probability structure (FQCP-structure) if and only if the following axioms are satisfied for all variables running over $\mathcal{C}$, provided that in the formula $A/B \precsim C/D$ the events B and D are elements of $\mathcal{C}_0 = \{A : A \in \mathcal{C} \ \& \ \emptyset/\Omega \prec A/\Omega\}$:

T_0 Ω is a nonempty finite set; $\mathcal{C}$ is the Boolean algebra of subsets of Ω, and $\precsim$ is a quaternary relation on $\mathcal{C}$;

T_1 $\emptyset/\Omega \prec \Omega/\Omega$;

T_2 $\emptyset/A \precsim B/C$;

T_3 $A/B \precsim AB/B$;

T_4 $A/B \precsim C/D \vee C/D \precsim A/B$;

T_5 $$\bigvee_{0<k\le n}\Big[A_k/\bigcap_{0\le i<k} A_i \precsim B_{\beta_k}/\bigcap_{0\le i<\beta k} B_i\Big] \Longrightarrow$$

$$\Longrightarrow \bigcap_{0<i\le n} A_i / A_0 \precsim \bigcap_{0<i\le n} B_{\beta_i} / B_0$$

for all permutations β on $\{1, 2, \ldots, n\}$; moreover, if in the antecedent $\prec$ holds for some k, then $\prec$ holds in the consequent;

T_6 $$\bigvee_{i<n}[A_i/B_i \precsim C_i/D_i] \Longrightarrow C_n/D_n \precsim A_n/B_n, \text{ if}$$

$$\sum_{i\le n} \hat{A}_i/B_i = \sum_{i\le n} \hat{C}_i/D_i .$$

Remarks:

(i) $A/B \succ C/D$ of course means $\neg\, C/D \precsim A/B$, and

$A/B \sim C/D$ means $A/B \precsim C/D$ & $C/D \precsim A/B$.

$\hat{A}/B$ denotes a vector in the quotient vector space $\mathcal{V}_B$; for details of $\mathcal{V}_B$ we refer the reader to Theorem 10. In fact, $\hat{A}/B$ can be identified with the partial characteristic function of A : $[\hat{A}/B](\omega) = 1$, if $\omega \in AB$, $[\hat{A}/B](\omega) = 0$, if $\omega \in \bar{A}B$, and is undefined elsewhere. Thus axiom T_6 can be stated purely in terms of elements of $\mathcal{A}$ and the relation $\precsim$. Note that the set $\mathcal{A} - \mathcal{A}_0$ is an ideal.

(ii) As one can see immediately, the crucial axioms are T_5 and T_6, corresponding to multiplication and addition laws of probability respectively. If we admit that $A/B \in \mathcal{A}/B$, then T_3 is trivially satisfied and $\precsim$ is a binary relation on the set of conditional events $\{A/B : A \in \mathcal{A}$ & $B \in \mathcal{A}_0\}$.

(iii) We are using the same symbols in several different senses in the present work. In particular, $< \Omega , \mathcal{A}, \precsim >$ has several different meanings in different contexts. This leads to typographical simplicity. There is, of course, always the possibility of designing some more ingenious symbolism.

(iv) The definition of an infinite qualitative conditional probability structure to be represented by a probability measure on $\mathcal{A}$ would be apparently quite messy and completely unintuitive. Some topological properties of $\precsim$ could make the formulation more agreeable. But this will be not our concern now.

Before we turn to some further details about the FQCP-structures, we should perhaps first examine the power of Definition 3 by listing its main consequences.

THEOREM 9 *Let* $<\Omega, \mathcal{A}, \precsim>$ *be a* FQCP-structure. *Then the following formulas are valid for all variables running over* $\mathcal{A}$ *provided that in* A/B , B *is restricted to* $\mathcal{A}_0$:

(1) $A/B \sim A/B$;

(2) $A/B \precsim C/D \;\&\; C/D \precsim E/F \Rightarrow A/B \precsim E/F$;

(3) $A/B \sim C/D \Rightarrow C/D \sim A/B$;

(4) $A/B \sim C/D \;\&\; C/D \sim E/F \Rightarrow A/B \sim E/F$;

(5) $A/B \precsim C/D \;\&\; C/D \prec E/F \Rightarrow A/B \prec E/F$;

(6) $A/B \prec C/D \;\&\; C/D \precsim E/F \Rightarrow A/B \prec E/F$;

(7) $A/B \prec C/D \;\&\; C/D \prec E/F \Rightarrow A/B \prec E/F$;

(8) $\sim$ *is an equivalence relation*;

(9) $A/B \prec C/D \vee C/D \prec A/B \vee A/B \sim C/D$ *and each of the formulas excludes the other two*;

(10) $A/C \precsim B/C \Leftrightarrow A \cup D/C \precsim B \cup D/C$, *if* $A, B \perp D$;

(11) $A/C \sim B/C \Leftrightarrow A \cup D/C \sim B \cup D/C$, *if* $A, B \perp D$;

(12) $A \subseteq B \Rightarrow A/C \precsim B/C$;

(13) $AB/B \precsim A/B$;

(14) $A/B \sim AB/B$;

(15) $A_1/C \precsim B_1/D \;\&\; A_2/C \precsim B_2/D \Rightarrow A_1 \cup A_2/C \precsim B_1 \cup B_2/D$, *if* $B_1 \perp B_2$;

(16) $A_1 \cup A_2 / C \precsim B_1 \cup B_2 / D \Rightarrow [A_1 / C \precsim B_1 / D \vee A_2 / C \precsim B_2 / C]$,

<u>if</u> $A_1 \perp A_2$;

(17) $\emptyset/B \sim \emptyset/\Omega$;

(18) $A/B \sim \emptyset/\Omega$, <u>if</u> $A \perp B$;

(19) (i) $A/B \precsim C/D \Leftrightarrow AB/B \precsim CD/D$;

(ii) $A/B \sim C/D \Leftrightarrow AB/B \sim CD/D$;

(iii) $A/B \prec C/D \Leftrightarrow AB/B \prec CD/D$;

(20) $A_1/B_1C_1 \precsim A_2/B_2C_2 \;\&\; B_1/C_1 \precsim B_2/C_2 \Rightarrow A_1B_1/C_1 \precsim A_2B_2/C_2$;

(21) $A_1/B_1C_1 \precsim B_2/C_2 \;\&\; B_1/C_1 \precsim A_2/B_2C_2 \Rightarrow A_1B_1/C_1 \precsim A_2B_2/C_2$;

(22) $A_1/B_1C_1 \precsim A_2/B_2C_2 \;\&\; B_1/C_1 \prec B_2/C_2 \Rightarrow A_1B_1/C_1 \prec A_2B_2/C_2$;

(23) $A_1/B_1C_1 \precsim B_2/C_2 \;\&\; B_1/C_1 \prec A_2/B_2C_2 \Rightarrow A_1B_1/C_1 \prec A_2B_2/C_2$;

(24) $A_1/B_1C_1 \prec A_2/B_2C_2 \;\&\; B_1/C_1 \precsim B_2/C_2 \Rightarrow A_1B_1/C_1 \prec A_2B_2/C_2$;

(25) $A_1/B_1C_1 \prec B_2/C_2 \;\&\; B_1/C_1 \precsim A_2/B_2C_2 \Rightarrow A_1B_1/C_1 \prec A_2B_2/C_2$;

(26) $A_1/B_1C_1 \sim A_2/B_2C_2 \;\&\; B_1/C_1 \sim B_2/C_2 \Rightarrow A_1B_1/C_1 \sim A_2B_2/C_2$;

(27) $A_1/B_1C_1 \sim B_2/C_2 \;\&\; B_1/C_1 \sim A_2/B_2C_2 \Rightarrow A_1B_1/C_1 \sim A_2B_2/C_2$;

(28) $\forall_{0 < k < n} [A_k / \bigcap_{0 \le i < k} A_i \precsim B_{\beta_k} / \bigcap_{0 \le i < \beta_k} B_i] \Longrightarrow$

$B_{\beta_n} / \bigcap_{0 \le i < \beta_n} B_i \precsim A_n / \bigcap_{0 \le i < n} A_i$, if

$\bigcap_{0 < i \le n} A_i / A_0 \sim \bigcap_{0 < i \le n} B_i / B_0$, <u>and</u> <u>if</u> <u>in</u> <u>the</u>

<u>antecedent</u> $\prec$ <u>holds</u> <u>for</u> <u>some</u> k , <u>then</u> <u>in</u> <u>the</u> <u>consequent</u>

$\prec$ <u>also</u> <u>holds</u>;

For notational convenience we introduce three vectors of relations Γ, Λ, Ξ, defined as follows: $\Gamma = < \precsim, \sim, \precsim, \prec, \prec >$, $\Lambda = < \precsim, \sim, \prec, \precsim, \prec >$, $\Xi = < \precsim, \sim, \prec, \prec, \prec >$, with coordinates Γ_i, Λ_i, Ξ_i, for $i = 1, 2, \ldots, 5$ respectively. In the following six clauses we assume $A_j \subseteq B_j \subseteq C_j$ for $j = 1, 2$; then for all $i = 1, 2, \ldots, 5$,

(29) $A_1/B_1 \; \Gamma_i \; A_2/B_2 \;\&\; B_1/C_1 \; \Lambda_i \; B_2/C_2 \Rightarrow A_1/C_1 \; \Xi_i \; A_2/C_2$;

(30) $A_1/B_1 \; \Gamma_i \; B_2/C_2 \;\&\; B_1/C_1 \; \Lambda_i \; A_2/B_2 \Rightarrow A_1/C_1 \; \Xi_i \; A_2/C_2$;

(31) $A_1/C_1 \; \Gamma_i \; A_2/C_2 \;\&\; B_2/C_2 \; \Lambda_i \; B_1/C_1 \Rightarrow A_1/B_1 \; \Xi_i \; A_2/B_2$;

(32) $A_1/C_1 \; \Gamma_i \; A_2/C_2 \;\&\; B_2/C_2 \; \Lambda_i \; A_1/B_1 \Rightarrow B_1/C_1 \; \Xi_i \; A_2/B_2$;

(33) $A_1/C_1 \; \Gamma_i \; A_2/C_2 \;\&\; A_2/B_2 \; \Lambda_i \; B_1/C_1 \Rightarrow A_1/B_1 \; \Xi_i \; B_2/C_2$;

(34) $A_1/C_1 \; \Gamma_i \; A_2/C_2 \;\&\; A_2/B_2 \; \Lambda_i \; A_1/B_1 \Rightarrow B_1/C_1 \; \Xi_i \; B_2/C_2$.

In the following we assume $A \subseteq B_1 \subseteq C \;\&\; A \subseteq B_2 \subseteq C$; furthermore, $:$ may denote any one of the following relations: $\precsim$, $\prec$, $\sim$. Then,

(35) $A/B_1 : A/B_2 \Leftrightarrow B_2/C : B_1/C$;

(36) $A/B_1 : B_2/C \Leftrightarrow A/B_2 : B_1/C$;

(37) $B_1/C : A/B_1 \Leftrightarrow B_2/C : A/B_1$;

(38) $AC/BD : AB/CD \Leftrightarrow C/D : B/D$;

(39) $AC/BD : C/D \Leftrightarrow AB/CD : B/D$;

(40) $B/D : AB/CD \Leftrightarrow C/D : AC/BD$;

(41) $A/A \sim B/B$;

(42) $\emptyset/A \prec B/B$;

(43) $\Omega/A \sim \Omega/\Omega$;

(44) $\Omega/A \sim \Omega/B$;

(45) $A/B \precsim C/D \Longleftrightarrow \overline{C}/D \precsim \overline{A}/B$;

(46) $A/B \precsim \Omega/\Omega$;

(47) $AB/\Omega \precsim A/B$;

(48) $AB/C \precsim A/BC$;

Again : *is to be read as any one of* $\precsim$, $\prec$, $\sim$. *Then,*

(49) $A_1B/\Omega : A_2B/\Omega \Longleftrightarrow A_1/B : A_2/B$;

(50) $A/BC : D/E \;\&\; A/B\overline{C} : D/E \Longrightarrow A/B : D/E$;

(51) $A/B : C/DE \;\&\; A/B : C/D\overline{E} \Longrightarrow A/B : C/D$;

(52) $A/B : A/BC \Longleftrightarrow A/B\overline{C} : A/B$;

(53) $\forall_{i < n}[A_i/B_i \prec C_i/D_i] \Longrightarrow C_n/D_n \prec A_n/B_n$, *if*

$$\sum_{i \leq n} \hat{A}_i/B_i = \sum_{i \leq n} \hat{C}_i/D_i \; ;$$

(54) $[A_1/A \precsim A_2/A \precsim \ldots \precsim A_n/A \;\&\; B_1/B \precsim B_2/B \precsim \ldots \precsim B_n/B] \Longrightarrow$

$\Longrightarrow A_1/A \precsim B_n/B$, *if* $A = \bigcup_{i=1}^{n} A_i$ & $B = \bigcup_{i=1}^{n} B_i$ & $A_i \perp A_j$,

$B_i \perp B_j$ *for* $i \neq j,\ 1 \leq i,\ j \leq n$;

(55) *If* $A \precsim^* B \Longleftrightarrow A/\Omega \precsim B/\Omega$, *then* $<\Omega , \mathcal{A}, \precsim^*>$ *is a* FQP-structure;

(56) $\forall_{0 \leq i < n}[A_i/A_{i+1} \precsim B_i/B_{\beta_i+1}] \longrightarrow B_{\beta_n}/B_{\beta_n+1} \precsim A_n/A_{n+1}$ *for all permutations* β *on* $\{1, 2, \ldots, n\}$, *where* $A_i \subseteq A_{i+1}$ $B_i \subseteq B_{i+1}$ $(i = 0, 1, \ldots, n)$, *and* $A_0/A_{n+1} \sim B_0/B_{n+1}$; *and if in the antecedent* $\prec$ *holds for some* k , *so does it in the consequent.*

Proof:

(1) Substitute in T_4 and use the definition of $\sim$.

(2) Since $\hat{A}/B + \hat{C}/D + \hat{E}/F = \hat{C}/D + \hat{E}/F + \hat{A}/B$, and (by assumption) $A/B \precsim C/D$ & $C/D \precsim E/F$, T_6 gives us $A/B \precsim E/F$.

(3) Use the definition of $\sim$.

(4) Use (2) twice.

(5) Obviously by (2) we have $A/B \precsim E/F$. If $A/B \sim E/F$ were the case for some $A, B, E,$ and $F,$ then $E/F \precsim A/B$ would be true, and hence by (2) $E/F \precsim C/D$ also, contrary to the assumption.

(6) Clearly $A/B \precsim E/F$. If $A/B \sim E/F$ were true for some $A, B, E,$ and $F,$ then also $E/F \precsim A/B$; Thus by (5) we get $E/F \prec C/D$, contrary to assumption.

(7) The assumption implies $A/B \prec C/D$ & $C/D \precsim E/F$; we can therefore use (6).

(8) Check (1), (3), and (4).

(9) Use T_4 and the definitions of $\prec$ and $\sim$.

(10) Since $\hat{A}/C + \hat{B}/C + \hat{D}/C = \hat{B}/C + \hat{A}/C + \hat{D}/C$ and $A, B \perp D$, we have $\hat{A}/C + (B \cup D)^\wedge/C = \hat{B}/C + (A \cup D)^\wedge/C$; so, using T_6, we get the equivalence.

(11) Use (10) twice.

(12) $A \subseteq B$ implies $B = A \cup B\bar{A}$ and also $\hat{B}/C + \hat{\emptyset}/C = \hat{A}/C + (B\bar{A})^{\wedge}/C$.

Since $\emptyset/C \precsim B\bar{A}/C$ by T_2, using T_6 we get $A/C \precsim B/C$.

(13) Since $AB \subseteq A$, we can use (12).

(14) Use (13) and T_3.

(15) If $B_1 \perp B_2$, $\hat{\emptyset}/C + \hat{A}_1/C + \hat{A}_2/C + (B_1 \cup B_2)^{\wedge}/D =$

$= \hat{B}_1/D + \hat{B}_2/D + (A_1 \cup A_2)^{\wedge}/C + (A_1A_2)^{\wedge}/C$; by T_2

$\emptyset/C \precsim A_1A_2/C$, and so, using the assumptions, we get the conclusion via T_6.

(16) (i) Suppose $A_1 \cup A_2/C \precsim B_1 \cup B_2/D$ and <u>not</u> $A_1/C \precsim B_1/D$.

Then by T_4 $B_1/D \prec A_1/C$. Since

$\hat{\emptyset}/C + (A_1 \cup A_2)^{\wedge}/C + \hat{B}_1/D + \hat{B}_2/D = \hat{A}_1/C + \hat{A}_2/C + (B_1 \cup B_2)^{\wedge}/D +$

$+ (B_1B_2)^{\wedge}/D$, T_6 gives $A_2/C \precsim B_2/D$.

(ii) Suppose $A_1 \cup A_2/C \precsim B_1 \cup B_2/D$ and <u>not</u> $A_2/C \precsim B_2/D$.

Then by T_4 we have $B_2/D \prec A_2/C$. As before, the assumptions and T_6 give $A_1/C \precsim B_1/D$.

(17) $\emptyset/B \precsim \emptyset/\Omega$ by T_2. But also $\emptyset/\Omega \precsim \emptyset/B$.

(18) $A/B \sim AB/B$ by (14). Since $A \perp B$, we have $A/B \sim \emptyset/B$.

Finally, using (17) and (4) we have $A/B \sim \emptyset/\Omega$.

(19) (i) $A/B \precsim C/D \Rightarrow AB/B \precsim A/B \precsim C/D \precsim CD/D$ (Use (13), T_3, and (2)). But also $AB/B \precsim CD/D \Rightarrow A/B \precsim AB/B \precsim CD/D \precsim C/D$.

(ii) Use case (i) twice.

(iii) Contrapositive of (i).

(20) Special case of T_5: put $n = 2$ and take β to be the identical permutation on $\{1, 2\}$.

(21) Special case of T_5: put $n = 2$ and take β to be the reversed permutation on $\{1, 2\}$.

(22) - (27) Special cases of T_5.

(28) Use T_5 and prove by contradiction.

(29) - (34) Special cases of (22) - (27).

(35) - (40) Proofs are analogous to those of (29) - (34).

(41) From (39) we get $CD/CD \preccurlyeq D/D$ and from (40) $D/D \preccurlyeq CD/CD$. Hence putting $D = \Omega$ we have $\Omega/\Omega \sim C/C$ for any C .

(42) $\emptyset/A \sim \emptyset/\Omega \prec \Omega/\Omega \sim B/B$, hence $\emptyset/A \prec B/B$.

(43) $\Omega/A \sim A/A \sim \Omega/\Omega$.

(44) $\Omega/A \sim A/A \sim B/B \sim \Omega/B$.

(45) Since $\hat{A}/B + \hat{\bar{A}}/B + \hat{\Omega}/D = \hat{C}/D + \hat{\bar{C}}/D + \hat{\Omega}/B$, we use T_6.

(46) $\emptyset/\Omega \preccurlyeq \bar{A}/B \Longleftrightarrow A/B \preccurlyeq \Omega/\Omega$.

(47) $AC/\Omega \preccurlyeq A/C \Longleftrightarrow C/\Omega \preccurlyeq \Omega/\Omega$ by (38).

(48) From (38) we get $AC/D \preccurlyeq A/CD \Longleftrightarrow C/D \preccurlyeq \Omega/D$, and $\Omega/D \sim \Omega/\Omega$. Hence (46) gives us the result.

(49) - (51) Use (35) - (40).

(52) $A/B \preccurlyeq A/BC \Longleftrightarrow C/D \preccurlyeq C/AB \Longleftrightarrow \bar{C}/AB \preccurlyeq \bar{C}/B \Longleftrightarrow A/B\bar{C} \preccurlyeq A/B$.

(53) If $A_n/B_n \preccurlyeq C_n/D_n$, let us put $E_i = A_{i+1}$, $F_i = B_{i+1}$, $G_i = C_{i+1}$, $H_i = D_{i+1}$ for $1 \leq i \leq n-1$ and $E_n = A_1$, $F_n = B_1$, $G_n = C_1$, $H_n = D_1$. Then from the assumption we get $\forall_{i < n}[E_i/F_i \preccurlyeq G_i/H_i]$, and hence by T_6 $G_n/H_n \preccurlyeq E_n/F_n$, which is impossible.

(54) Assume $B_n/B \prec A_1/A$. Then $B_i/B \preceq B_n/B \prec A_1/A \preceq A_i/A$ $\Longrightarrow B_i/B \prec A_i/A$ for all $i = 1, 2, \ldots, n$. Since $\hat{B}_1/B + \ldots + \hat{B}_n/B + \hat{A}/A = \hat{A}_1/A + \ldots + \hat{A}_n/A + \hat{B}/B$, and $B_i/B \prec A_i/A$, by (53) we have $B/B \prec A/A$ which is impossible.

(55) Axioms T_1, T_2, T_4, and T_6 reduce to Scott's axioms for FQP-structures, if we put Ω for B in all terms of the form A/B.

(56) Trivial consequence of (28). Q. E. D.

Notice that Theorem 9 is also a consequence of Definition 2 and Theorem 6, if we put $A/B \preceq C/D$ equivalent to $AB \times D \preceq CD \times B$. On the other hand, if we let $A \parallel B$ mean $A/B \sim A/\Omega$, then Theorem 8 becomes a consequence of Definition 3 and Theorem 9. This interplay goes further. We can put $A\ C_+\ B \Leftrightarrow A/\Omega \prec A/B$ and $A\ C_-\ B \Leftrightarrow A/B \prec A/\Omega$, and also $A/C \parallel B/C \Leftrightarrow A/C \sim A/BC$; thence we can derive the basic properties of these notions in qualitative terms. Again, we can put $A \preceq B \Leftrightarrow A/\Omega \preceq B/\Omega$, $A \preceq B/C \Leftrightarrow A/\Omega \preceq B/C$ and $A/B \preceq C \Leftrightarrow A/B \preceq C/\Omega$, and handle the qualitative (absolute) probability relation as a special case of qualitative conditional probability relation.

Let $< \Omega, \mathcal{Cl}, P >$ be a finite probability space and let $\mathcal{P}$ be a partition of Ω. Then the function $P(A/\mathcal{P}) = \sum_{B \in \mathcal{P}} \hat{A} \cdot P(A/B)$ is called the global conditional probability measure of the event A, given the experiment (partition) $\mathcal{P}$. Note that the value of this measure is a function and not a real number, and that the following are true:

(i) $0 \leq P(A/\mathscr{P}) \leq 1$,

(ii) $P(A \cup B/\mathscr{P}) = P(A/\mathscr{P}) + P(B/\mathscr{P})$, if $A \perp B$,

(iii) $P(A/\mathscr{P}) = \hat{A}$, if $A \in \mathscr{P}$,

(iv) $P(A/\mathscr{P}) = \hat{\Omega} \cdot P(A)$, if $\forall_B[B \in \mathscr{P} \Longrightarrow B \parallel A]$, where

$A, B \in \mathscr{U}$, and $\mathscr{P}$ is a partition of Ω .

One might wonder if there is such an entity as a <u>globally conditionalized</u> event: $A/\mathscr{P}$. Such 'events' would be particularly interesting because we know that iteration of conditionalizations by events $(\ldots((A_0/A_1) / A_2) / \ldots) / A_n$ does not lead to anything new, since this is equal to $A_0 / \bigcap_{i=1}^{n} A_i$. But we might hope to get some new entities by changing the conditionalizing entities. We know that the Boolean closure $\mathscr{U}[\mathscr{P}]$ of $\mathscr{P}$ is a Boolean subalgebra of $\mathscr{U}$; and, vice versa, any Boolean subalgebra $\mathscr{B}$ of $\mathscr{U}$ defines exactly one partition $\mathscr{P}$ of Ω , $\mathscr{P}$ being just the set of atoms of $\mathscr{B}$. (Remember that we are working now with finite Boolean algebras.) Therefore it seems reasonable to consider $A/\mathscr{P}$ as an element of the quotient Boolean algebra $\mathscr{U}/\mathscr{U}[\mathscr{P}]$, where analogously to the case of A/B (where we relativized the set of possible outcomes to B), we now relativize the set of possible events to the Boolean algebra $\mathscr{U}[\mathscr{P}]$. The symbol $A/\mathscr{P}$ then becomes a legitimate set-theoretic entity, with a clear probabilistic meaning:

$A/\mathscr{P}$ = the set of events <u>indistinguishable</u> from the event A, given the events in the aglebra $\mathscr{U}[\mathscr{P}]$, generated by the experiment $\mathscr{P}$.

We shall come back to this problem in Section 3.3.

The notion of a globally conditionalized event plays an important role in advanced probability theory, and it may be of some methodological interest to study a qualitative probability relation on these entities. But beyond stating the problem, we shall not dig deeper into the matter here.

We now turn to the representation theorem for FQCP-structures.

THEOREM 10 Let $< \Omega , \mathcal{A}, \preccurlyeq >$ be a finite structure, where Ω is a nonempty finite set; $\mathcal{A}$ is the Boolean algebra of subsets of Ω , and $\preccurlyeq$ is a quaternary relation on $\mathcal{A}$. Let

$$\mathcal{A}_0 = \{A : \emptyset/\Omega \prec A/\Omega\} .$$

Then $< \Omega , \mathcal{A}, \preccurlyeq >$ is a FQCP-structure if and only if there exists a finitely additive conditional probability measure on $\mathcal{A}$ such that $< \Omega , \mathcal{A}, \mathcal{A}_0 , P >$ is a conditional probability space, and for all $A, C \in \mathcal{A}$ and $B, D \in \mathcal{A}_0$:

$$A/B \preccurlyeq C/D \Leftrightarrow P(A/B) \leq P(C/D) .$$

Proof:

I. The existence of a conditional probability measure on $\mathcal{A}$.

Suppose that $< \Omega , \mathcal{A}, \preccurlyeq >$ is a FQCP-structure. Let us define m real n-dimensional vector spaces $\mathcal{V}_B$ ($m = |\mathcal{A}_0|$, $n = |\Omega|$, $B \in \mathcal{A}_0$) as follows: The basis of $\mathcal{V}_B$ is the set $\{<(\{\omega\})^{\wedge}, B >\}_{\omega \in \Omega}$, where as usual, the hat '^' denotes the characteristic function of the given set, written in the form of an n-dimensional vector:

$$\hat{A} = < \hat{A}(\omega_1), \hat{A}(\omega_2), \ldots, \hat{A}(\omega_n) >$$

where $\Omega = \{\omega_i\}_{i=1}^{n}$. In particular,

$< (A \cup B)\hat{}, C > = < \hat{A}, C > + < \hat{B}, C >$ for $A \perp B$, and

$\alpha < \hat{A}, C > = < \alpha\hat{A}, C >$ for $A, B \in \mathcal{U}$ and $C \in \mathcal{U}_0$.

In fact, in the ordered couple $< \hat{A}, C >$, C is just an index from $\mathcal{U}_0$. We put $\hat{A}/C$ for $< \hat{A}, C >$, in order to simplify the notation.

If we take the (external) direct sum $\mathcal{W} = \bigoplus_{A \in \mathcal{U}_0} \mathcal{V}_A$ of all indexed vector spaces $\mathcal{V}_A$ for $A \in \mathcal{U}_0$, then the vectors in $\mathcal{W}$ are m-tuples

$$< v_1/A_1, v_2/A_2, \ldots, v_m/A_m > , \text{ where } \{A_i\}_{i=1}^{m} = \mathcal{U}_0$$

and $v_i \in \mathcal{V}(\Omega)$, for $i = 1, 2, \ldots, m$. The operations in $\mathcal{W}$ satisfy:

(i) $< v_1/A_1, v_2/A_2, \ldots, v_m/A_m > + < w_1/A_1, w_2/A_2, \ldots, w_m/A_m > =$

$= < v_1 + w_1/A_1, v_2 + w_2/A_2, \ldots, v_m + w_m/A_m >$;

(ii) $\alpha < v_1/A_1, v_2/A_2, \ldots, v_m/A_m > = < \alpha v_1/A_1, \alpha v_2/A_2, \ldots, \alpha v_m/A_m >$;

where $v_i, w_i \in \mathcal{V}(\Omega)$ for $i = 1, 2, \ldots, m$, and $\alpha \in Re$.

Obviously $\mathcal{V}_{A_i} \cong \mathcal{W}_{A_i}$, if $\mathcal{W}_{A_i}$ is the subspace of vectors of $\mathcal{W}$ of the form

$$< 0/A_1,\ 0/A_2,\ \ldots,\ 0/A_{i-1},\ v/A_i,\ 0/A_{i+1},\ \ldots,\ 0/A_m > ,$$

where $v \in \mathcal{V}(\Omega)$, $i = 1, 2, \ldots, m$.

We can in a one-one way associate with the entity A/B a vector $< 0/A_1, \ldots, \hat{A}/B, \ldots, 0/A_m >$ from $\mathcal{W}_B$, and put

$$A/B \preccurlyeq C/D \Leftrightarrow < 0/A_1, \ldots, \hat{A}/B, \ldots, 0/A_m > \preccurlyeq < 0/A_1, \ldots, \hat{C}/D, \ldots, 0/A_m > ,$$

so that the problem is finally formulated in the geometric language of vector spaces. In particular, the set of conditional entities $\{A/B : A \in \mathcal{Cl} \ \&\ B \in \mathcal{Cl}_0\}$ is 'translated' into a nonempty finite set of vectors from $\mathcal{W}$ with rational coordinates with respect to the basis

$$\{< 0/A_1, \ldots, 0/A_{i-1}, (\{\omega\})\hat{}/A_i, 0/A_{i+1}, \ldots, 0/A_m > : \omega \in \Omega \ \&\ i=1,2,\ldots, m\}.$$

Having done this, we are ready to use Corollary 5. In fact, translating T_4 and T_6 into vector language, we get the necessary and sufficient conditions for the existence of a linear functional $\psi : \mathcal{W} \longrightarrow \mathrm{Re}$ such that

$$\hat{A}/B \preccurlyeq \hat{C}/D \Leftrightarrow \psi(\hat{A}/B) \leq \psi(\hat{C}/D)$$

is satisfied for all $A, C \in \mathcal{Cl}$, $B, D \in \mathcal{Cl}_0$; here we put $\hat{A}/B$ for $< 0/A_1, \ldots, \hat{A}/B, \ldots, 0/A_m >$ and $\hat{C}/D$ for

$< 0/A_1, \ldots, \hat{C}/D, \ldots, 0/A_m >$. T_2 implies $\psi(\hat{A}/B) \geq \psi(\hat{\emptyset}/C) = 0$

and T_1 forces ψ to be strictly positive for $\hat{\Omega}/\Omega$. In particular, ψ can be normalized by defining

$$\varphi(\hat{A}/B) = \frac{\psi(\hat{A}/B)}{\psi(\hat{\Omega}/\Omega)} .$$

Theorem 9(14) gives us $\varphi(\hat{A}/B) = \varphi((AB)^\wedge/B)$. Suppose $A \subseteq B$ and $C \subseteq D$; then putting $\hat{A} \otimes \hat{D} \precsim \hat{C} \otimes \hat{B} \Leftrightarrow \hat{A}/B \precsim \hat{C}/D$, we can translate the countably many consequences of T_5 into consequences of Q_5 (Definition 2 in Section 2.4). Then, as in the case of the representation of FAQQP-structures, we apply the theory of nets (Y. Aczél, G. Pickert, and F. Radó [35]). In particular, $\hat{A}/A \sim \hat{B}/B$ is translated into $\hat{A} \otimes \hat{B} \sim \hat{B} \otimes \hat{A}$; if $A_i \subseteq B_i \subseteq C_i$ for $i = 1, 2$, then $\hat{A}_1/B_1 \precsim \hat{A}_2/B_2$ & $\hat{B}_1/C_1 \precsim \hat{B}_2/C_2 \Rightarrow \hat{A}_1/C_1 \precsim \hat{A}_2/C_2$ is translated into $\hat{A}_1 \times \hat{B}_2 \precsim \hat{A}_2 \times \hat{B}_1$ & $\hat{B}_1 \times \hat{C}_2 \precsim \hat{B}_2 \times \hat{C}_1 \Rightarrow \hat{A}_1 \times \hat{C}_2 \precsim \hat{A}_2 \times \hat{C}_1$. Hence, as for FAQQP-structures, there must exist a linear functional $f : \mathcal{V}(\Omega) \longrightarrow Re$ (see the construction in the proof of Theorem 7 in Section 2.4) such that

$$\varphi(\hat{A}/B) \leq \varphi(\hat{C}/D) \Longleftrightarrow f(\hat{A}) \cdot f(\hat{D}) \leq f(\hat{C}) \cdot f(\hat{B})$$

for all $A, C \in \mathcal{U}$, $B, D \in \mathcal{U}_0$ such that $A \subseteq B$, $C \subseteq D$. Hence for some $\eta : [0, 1] \longrightarrow [0, 1]$, $\eta(\varphi(\hat{A}/B) = \dfrac{f(\hat{A})}{f(\hat{B})}$, if $A \subseteq B$. By the additivity of φ on $\mathcal{W}$ and of f on $\mathcal{V}(\Omega)$ we find η to be a constant mapping; and after normalization of f it becomes even the identity mapping. Thus, for $A \subseteq B \subseteq C$ we have $\varphi(\hat{A}/C) = \varphi(\hat{A}/B) \cdot \varphi(\hat{B}/C)$.

We can now collect the results of our proof in the following conditions:

(i) $0 \leq \varphi(\hat{A}/B) \leq 1$;

(ii) $\varphi(\hat{\Omega}/\Omega) = 1$;

(iii) $\varphi((A \cup B)^\wedge/C) = \varphi(\hat{A}/C) + \varphi(\hat{B}/C)$, if $A \perp B$;

(iv) $\varphi(\hat{A}/B) = \varphi((AB)^\wedge/B)$;

(v) $\varphi(\hat{A}/C) = \varphi(\hat{A}/B) \cdot \varphi(\hat{B}/C)$, if $A \subseteq B \subseteq C$.

It is easy to show that (iv) & (v) imply

$$\varphi((AB)^\wedge/C) = \varphi(\hat{A}/BC) \cdot \varphi(\hat{B}/C) .$$

Finally, if we put $P(A/B) = \varphi(\hat{A}/B)$ for $A \in \mathcal{A}$, $B \in \mathcal{A}_0$, we get the desired conditional probability measure for which

$$A/B \precsim C/D \Leftrightarrow P(A/B) \leq P(C/D) \text{ for all } A, C \in \mathcal{A} , B, D \in \mathcal{A}_0 . \quad (2.14)$$

II. Necessity.

It is a routine matter to check that the conditions $T_0 - T_6$ are also necessary for the existence of a conditional probability measure P on $\mathcal{A}$. Q. E. D.

One of the basic questions of Representation Theory is the problem of the uniqueness of the representing function. That is to say, we would like to know the structure of the class of measures P satisfying the representation condition (2.14).

Unfortunately in no structure here studied is the answer very simple. For example, in the case of the finite qualitative probability structures (FQP-structures), we can think of several apparently unrelated measures that represent the ordering $\precsim$. Given one measure, we can construct another by, roughly speaking, moving its values a little bit, keeping the additivity law valid, and at the same time not violating the validity of the inequality.

This construction can be made in uncountably many ways showing no particular structure. Similar problems will appear with qualitative information and entropy structures.

It is known that in atomless Boolean algebras the representing measure is unique. On the other hand, atomless Boolean algebras are not the most important ones.

The problem of uniqueness in general nonlinear measurement structures certainly deserves some further study.

Another problem is to find those conditions that must be imposed on the structure $\langle \Omega , \mathfrak{A} , \preccurlyeq , \| \rangle$ in order to find a probability measure P on $\mathfrak{A}$ such that

(i) $\quad A \preccurlyeq B \Leftrightarrow P(A) \leq P(B)$;

(ii) $\quad A \| B \Leftrightarrow P(AB) = P(A) \cdot P(B)$, where $A, B \in \mathfrak{A}$.

Yet another class of questions arises, when we want to represent the various possible combinations of the 'relations' $A \precsim B$, $A \| B$, $A/B \precsim C/D$, $A/C \| B/C$, $A \; C_+ \; B/C$, $A/C \; C_- \; B/C$, etc., by an adequate probability measure.

These problems are outside of the scope of this work.

2.7. Additively Semiordered Qualitative Conditional Probability Structures

In this section the solution of problem (P_4) will be discussed. In particular, the basic properties of the quaternary relation $\succ$ will be presented. The intended interpretation of the formula $A/B \succ C/D$ will be: event A given event B is definitely more probable than event C given event D. If we put $A/B \succcurlyeq C/D \Longleftrightarrow$ $\Longleftrightarrow [A/B \succ C/D \vee \neg\, C/D \succ A/B]$, then an equivalent structure is obtained.

Since (P_4) is a generalization of (P_3) and (P_2), one may expect that the properties of $\succ$ will resemble somewhat the properties of $\precsim$ in problem (P_3) and of $\succ$ in problem (P_2). In addition, the proof of the representation theorem will be a combination of proof techniques used for representation theorems of semiordered and conditional probability structures.

The relation $\succ$, called the <u>semiordered qualitative conditional probability relation</u> is at least a semiorder. The additional properties are dictated by the probabilistic interpretation.

Perhaps we should point out that some of the notions discussed in Section 2.6 have their 'semiorder' counterparts. For instance, $A\, C_+^s\, B \Longleftrightarrow A/B \succ A/\Omega$, $A\, C_-^s\, B \Longleftrightarrow A/\Omega \succ A/B$ ($A \in \mathscr{A}$ and $B \in \mathscr{A}_0$) are the <u>semiordered relevance relations</u>.

We shall not try to speculate about the use of these notions. Maybe they will have some importance in a rather general qualitative theory of causality.

Much recent work in inductive logic and methodology of science has been concerned with rules of <u>acceptance</u> or <u>rejection</u> (of scientific theories). The simplest rule studied allows a hypothesis or theory, represented by the event A, to be accepted iff $P(A) > 1 - \varepsilon$ $(0 < \varepsilon \leq 1/2)$, or in the conditional version, $P(A/B) > 1 - \varepsilon$.

In terms of FASQP-structures or semiordered qualitative conditional probability structures this rule gets a more polished and symmetric form:

(i) A is accepted $\Leftrightarrow A \approx \Omega$;

A is rejected $\Leftrightarrow A \approx \emptyset$;

(ii) A is accepted given $B \Leftrightarrow A/B \approx \Omega/\Omega$;

A is rejected given $B \Leftrightarrow A/B \approx \emptyset/\Omega$.

Since $\approx$ is not transitive, one cannot hope to describe too much with it. The set of all accepted events does not form even a filter (it is true that $A \approx \Omega \ \& \ A \subseteq B \Rightarrow B \approx \Omega$, but $A \approx \Omega \ \& \ B \approx \Omega \Rightarrow AB \approx \Omega$ is false) which is an obvious algebraic requirement of a deductive system! The only possible way to remove this weakness while retaining the rule, is to think of $\mathcal{U}$ as a kind of lattice rather than a Boolean algebra; in this lattice $\approx$ will still have the probabilistic interpretation, but the representation theorem for this new structure may fail. After all, the acceptance and rejection predicates are supposed to be meant for (empirical) theories; and the set of these forms at best a Brouwerian algebra. Moreover, we hardly would want to consider numerical probabilities for evaluating the degree of acceptance, since we defined the notions of acceptance and rejection in terms of inequalities. Thus, considering $\mathcal{U}$ as a lattice and $\approx$ as a probabilistic indifference relation, we may conceivably get a deductive system of accepted theories. But we shall not deal further with this rather delicate philosophical problem. In general, threshold-type statements are always rather weak from the point of view of their content.

We turn now to the definition of FASQCP-structures.

DEFINITION 4 *A triple* $< \Omega , \mathfrak{A}, \succ >$ *is a finitely additive semiordered qualitative conditional probability structure* (FASQCP-structure) *iff the following axioms are satisfied when all variables run over* $\mathfrak{A}$; *provided that, in the formula* $A/B \succ C/D$, *the events* B *and* D *are elements of* $\mathfrak{A}_0 = \{A \in \mathfrak{A} : \neg\, \emptyset/\Omega \sim A/\Omega\}$:

R_0 Ω *is a nonempty finite set;* $\mathfrak{A}$ *is the Boolean algebra of subsets of* Ω , *and* $\succ$ *is a quaternary relation on* $\mathfrak{A}$;

R_1 $\Omega/\Omega \succ \emptyset/\Omega$;

R_2 $\neg\, A/B \succ A/B$;

R_3 $A/B \sim AB/B$;

R_4 $C/D \succ B/E \Longrightarrow C/D \succ A/E$, if $A \subseteq B$;

R_5 $$\forall_{0 < k \le n}\Big[A_k \Big/ \bigcap_{0 \le i < k} A_i \succ B_{\beta_k} \Big/ \bigcap_{0 \le i < \beta_k} B_i\Big] \Longrightarrow$$

$$\Longrightarrow \bigcap_{0 \le i \le n} A_i/A_0 \succ \bigcap_{0 \le i \le n} B_i/B_0$$ *for all permutations* β

on $\{1, 2, \ldots, n\}$, *if* $\neg\, \bar{B}_{k+1} \Big/ \bigcap_{0 \le i \le k} B_i \succ \bigcap_{0 \le i \le k} B_i/B_0$

for $k = 1, 2, \ldots, n-1$;

$R_6 \quad \forall_{i<n}[A_i/B_i \succ C_i/D_i \;\&\; \neg\, E_i/F_i \succ G_i/H_i] \Longrightarrow$

$\Longrightarrow [A_n/B_n \succ C_n/D_n \Rightarrow E_n/F_n \succ G_n/H_n]$ <u>if</u>

$\sum_{i \le n} [\hat{A}_i/B_i + \hat{G}_i/H_i] = \sum_{i \le n} [\hat{C}_i/D_i + \hat{E}_i/F_i]$;

Remarks:

(i) $A/B \sim C/D$ is of course equivalent to

$\forall_{E,F}[A/B \approx E/F \Leftrightarrow C/D \approx E/F]$, where $A/B \approx C/D$

means $\neg\, A/B \succ C/D \;\&\; \neg\, C/D \succ A/B$; several other notions can be introduced as in the case of FASQP-structures.

(ii) The assumption $\neg\, \bar{B}_{k+1} / \bigcap_{0 \le i \le k} B_i \succ \bigcap_{0 \le i \le k} B_i / B_0$ (2.15)

for $k = 1, 2, \ldots, n-1$ in axiom R_5 is a little bit strong, since $\mathcal{E} + P(B_i/C_i) \ge P(\bar{A}_i/B_iC_i)$ is enough, too. Because there is no way of representing a formula $\mathcal{E} + P(A/B) = P(C/D)$ in terms of $\succ$, we have to leave out the case of equality.

(iii) Axioms R_1 - R_6 are just the combinations of axioms for FASQP-structures and FQCP-structures. Certain axioms are given also by Suppes [16]. As expected, axiom R_6 is the qualitative version of the addition law, whereas R_5 provides the multiplication law. Naturally, all important properties of the relation $\succ$ are hidden in these two axioms.

THEOREM 11 *Let* $<\Omega, \mathcal{A}, \succ>$ *be a* FASQCP-structure. *Then the following formulas are valid for all variables running over* $\mathcal{A}$, *provided that in* A/B, B *is restricted to* $\mathcal{A}_0$:

(1) $A_1/B_1 \succ C_1/D_1 \ \& \ A_2/B_2 \succ C_2/D_2 \Rightarrow [A_1/B_1 \succ C_2/D_2 \vee A_2/B_2 \succ C_1/D_1]$;

(2) $A_1/B_1 \succ C/D \ \& \ C/D \succ E/F \Rightarrow [A_1/B_1 \succ A_2/B_2 \vee A_2/B_2 \succ E/F]$;

(3) $A/B \succ C/D \ \& \ E/B \succ B/D \Rightarrow A \cup E/B \succ C \cup G/D$, *if* $A \perp E$;

(4) $A \cup E/B \succ C \cup G/D \Rightarrow [A/B \succ C/D \quad E/B \succ G/D]$, *if* $C \perp G$;

(5) $A/B \succ C/D \Leftrightarrow \overline{C}/D \succ \overline{A}/B$;

(6) $A \subseteq B \Rightarrow \neg A/C \succ B/C$;

(7) $A/B \succ C/D \ \& \ C/D \succ E/F \Rightarrow A/B \succ E/F$;

(8) $\neg A/B \approx \emptyset/C \Rightarrow A/B \succ \emptyset/C$;

(9) $\forall_{i<n}[A_i/B_i \succ C_i/D_i] \rightarrow C_n/D_n \succ A_n/B_n$, *if*

$$\sum_{i \leq n} \hat{A}_i/B_i = \sum_{i \leq n} \hat{C}_i/D_i ;$$

(10) $A/B \succ C/D \Leftrightarrow \Omega/\Omega \succ C/D$, *if* $B \subseteq A$;

(11) $A/A \succ \emptyset/B$;

(12) $C/D \succ A/B \Rightarrow C/D \succ \emptyset/F$;

(13) $A_1/B_1C_1 \succ A_2/B_2C_2 \ \& \ B_1/C_1 \succ B_2/C_2 \Rightarrow A_1B_1/C_1 \succ A_2B_2/C_2$, *if*

$\neg A_2/B_2C_2 \succ B_2/C_2$;

(14) $A_1/B_1C_1 \succ B_2/C_2 \ \& \ B_1/C_1 \succ A_2/B_2C_2 \Rightarrow A_1B_1/C_1 \succ A_2B_2/C_2$,

$\underline{\text{if}} \quad \neg \bar{A}_2/B_2C_2 \succ B_2/C_2$;

(15) $A/B \succ C/D \Rightarrow A \cup E/B \succ C \cup F/D$, $\underline{\text{if}} \quad F \subseteq E \ \& \ E \perp A$;

The proof goes along the same lines as the proof of Theorems 4 and 9 above.

Note that all 'addition laws' go through smoothly (remember that $< \mathcal{U}, \succ >$ is a finitely additive semiordered structure), whereas the 'multiplication laws' sometimes fail. For instance, there is no simple counterpart of the theorem

$$A_1/C_1 \sim A_2/C_2 \Rightarrow [A_1/B_1 \precsim A_2/B_2 \Leftrightarrow B_2/C_2 \precsim B_1/C_1] ,$$

if $A_i \subseteq B_i \subseteq C_i \quad (i = 1, 2)$, which is valid for qualitative conditional probability structures. If $A \times B \succ C \times D$ denotes the semiorder version of the quadratic qualitative probability relation, then, as one can check easily, the transformation

$$A_1/B_1 \succ A_2/B_2 \rightsquigarrow A_1 \times B_1 \succ A_2 \times B_2 \quad \text{for} \quad A_i \subseteq B_i \quad (i = 1, 2)$$

is valid, but not conversely! Therefore we cannot hope to give a representation theorem in a complete form. The inequality $P(A/B) \geq P(C/D) + \mathcal{E} \quad (0 < \mathcal{E} \leq 1)$ behaves with respect to multiplication quite irregularly. For example, the standard cancellation law: $A \times B \precsim C \times D \ \& \ C \times E \precsim F \times B \Rightarrow A \times E \precsim F \times D$ is valid only under very special conditions.

More specifically, we are able to show the following theorem:

THEOREM 12 (Representation theorem) *Let* $< \Omega , \mathcal{A} , \succ >$ *be a finite structure, where* Ω *is a nonempty finite set;* $\mathcal{A}$ *is the Boolean algebra of subsets of* Ω *and* $\succ$ *is a quaternary relation on* $\mathcal{A}$; *let* $\mathcal{A}_0 = \{A : \neg\, \emptyset/\Omega \sim A/\Omega\}$.
Then $< \Omega , \mathcal{A} , \succ >$ *is a* FASQCP-structure if and only if *there exists a finitely additive probability measure* P *on* $\mathcal{A}$ *and a real number* ε *such that all variables run over* $\mathcal{A}$ *and the event* B *in* A/B *is restricted to* $\mathcal{A}_0$, *the following conditions are satisfied:*

(1) $A/B \succ C/D \quad P(A/B) \geq P(C/D) + \varepsilon$ *and* $0 < \varepsilon \leq 1$;

(2) $P(\Omega/\Omega) = 1$;

(3) $0 \leq P(A/B) \leq 1$

(4) $P(A \cup B/C) = P(A/C) + P(B/C)$, *if* $A \perp B$;

(5) $P(AB/B) = P(A/B)$;

(6) $$\forall_{0 < k \leq n} \Big[P\Big(A_k \Big/ \bigcap_{0 \leq i < k} A_i\Big) \geq P\Big(B_{\beta_k} \Big/ \bigcap_{0 \leq i \leq \beta_k} B_i\Big) + \varepsilon \Big]$$

$$\Longrightarrow P\Big(\bigcap_{0 \leq i \leq n} A_i \Big/ A_0\Big) \geq P\Big(\bigcap_{0 \leq i \leq n} B_{\beta_i} \Big/ B_0\Big) + \varepsilon$$

for all permutations β *on* $\{1, 2, \ldots, n\}$, *if*

$$P\Big(\overline{B}_{k+1} \Big/ \bigcap_{0 \leq i \leq k} B_i\Big) < P\Big(\bigcap_{0 \leq i \leq k} B_i \Big/ B_0\Big) + \varepsilon$$

for all $k = 1, 2, \ldots, n-1$.

The proof is a combination of the proofs of Theorems 5 and 10.

From a measurement-theoretic viewpoint FASQCP-structures are quite complicated. We can say nothing about the uniqueness of the representing measure, except that some periodic transformations (with period ε) should lead to new measures, satisfying conditions (1) - (6) in Theorem 12.

In closing this chapter, we can claim that the methodology described in Section 1.4 has turned out to be very useful in proving all the basic measurement-theoretic theorems about probabilistic relational structures. In the next two chapters we shall present some further applications of this method.

3. APPLICATIONS TO INFORMATION AND ENTROPY STRUCTURES

3.1. Recent Developments in Axiomatic Information Theory

Information theory deals with the mathematical properties of communication models, which are usually defined in terms of concepts like channel, source, information, entropy, capacity, code, and which satisfy certain conditions and axioms.

Our knowledge in this field has expanded prodigiously since C. E. Shannon gave in 1948 the first sufficiently general definition of information and entropy. An indication of this expansion can be gained from the survey and extensive bibliography in R. S. Varma and P. Nath [38]. In particular, the last ten years have seen a considerable interest in the abstract axiomatic treatment of the concepts of information and entropy.

Shannon's original axioms for the entropy measure have been replaced several times subsequently by weaker conditions (see Fadeev [39], Khinchin [40], Tveberg [41], Kendall [42], and others). The weakest set of axioms known seems to be that given by Lee [43].

Rényi [44], on the other hand, has extended the notion of entropy by using the concept of a generalized probability distribution.

The above characterizations of entropy all involve essentially probabilistic notions.

Ingarden and Urbanik [45, 46, 47], and de Fériet and Forte [48], have given axiomatic definitions of entropy and information measures without using probability measures. Similarly Kolmogorov [49, 50] has shown that the basic information-theoretic concepts can be formulated without recorse to probability theory.

Ingarden and Urbanik need to assume for their definition of entropy a sufficiently large pseudometric space of finite Boolean rings, in order to be able to state the continuity of the entropy measure. On the other hand, Kolmogorov uses the concepts of recursive function and random sequence. Still another approach is known in coding theory.

Quite recently, several information-theorists have tried to construct the information-theoretic notions by using techniques from statistical decision theory. For example, Belis and Guiasu [51] propose a notion of a 'qualitative-quantitative information measure,' defined in terms of utility. The idea is simply the following: Given a probability space $< \Omega , \mathcal{U}, P >$, they introduce, besides the probability measure P on the algebra of events $\mathcal{O}\!\mathcal{L}$, a utility function U , which assigns to each element of a partition $\mathcal{P}$ of Ω a non-negative real number: the entropy measure H of the partition $\mathcal{P}$ is then given by

$$H(\mathcal{P}) = - \sum_{A\in} U(A) \cdot P(A) \cdot \log_2 P(A) .$$

Weiss [52] gives an axiomatic system for subjective information which is almost identical with the theories of probability and utility of Savage [6] and Pratt, Raiffa and Schlaifer [53].

In a related field, that of semantic information theory (in the sense of Bar-Hillel and Carnap [54]), there have also been advances (see especially Hintikka [55, 56]).

As can be seen even from this cursory review of recent developments, there is available an immense wealth of axiomatic material dealing with purely logical and foundational aspects of information theory. The above-mentioned foundational attempts are all directed in the main towards axiomatizing the basic information-theoretic notions in the form of <u>functional equations</u>. In this paper another approach is proposed. We shall advocate, instead of the <u>analytic approach</u>, an <u>algebraic approach</u> in terms of <u>relational structures</u>. The latter approach is more relevant to <u>measurement</u> or, generally, <u>epistemic aspects</u> of information, unlike the former which tackles the <u>a priori</u>, or <u>ontological aspects</u> of information-theoretic problems.

In fact, the main purpose of this chapter is to give axiomatic definitions of the concepts of qualitative information and qualitative entropy structure, and to study some of their basic properties. The chapter culminates in proving certain representation theorems which elucidate the relations these notions bear to the standard concepts of information and entropy.

3.2. Motivations for Basic Notions of Information Theory

The standard notion of information is introduced usually in order to answer the following somewhat abstract question: How much information do we get about a point $\omega \in \Omega$ from the news that ω belongs to a subset A of Ω, that is $\omega \in A$ and $A \subseteq \Omega$? It is rather natural to assume that the answer should depend on, and only on, the size of A, that is to say, on P(A), where P is a standard probability measure on the Boolean algebra $\mathfrak{A}$ of subsets of Ω. In other words, the answer should be given in terms of a real-valued function I, defined on the unit interval [0, 1]. Hence, the amount of information conveyed by the statement $\omega \in A$ will be $I \circ P(A)$, or in a simpler notation, $I_P(A)$. It is also natural to require I to be non-negative and continuous on [0, 1]. Now, if we are given two independent experiments which are described by statements $\omega \in A$ and $\omega \in B$ $(A, B \in \mathfrak{A}, \omega \in \Omega)$, then it is reasonable to expect that the amount of information of the experiment described by $\omega \in A \;\&\; \omega \in B$, that is $\omega \in A \cap B$, will be the sum of the amounts of information of the experiments taken separately.

Given a probability space $\mathbb{A} = \langle \Omega, \mathcal{A}, P \rangle$, let $A \parallel B$ mean that the experiments with outcomes $\omega \in A$ and $\omega \in B$ are probabilistically independent $(A, B \in \mathcal{A})$; then we can collect our previous ideas in the following assumptions:

(i) The diagram

$$
\begin{tikzcd}
\mathcal{A} \arrow[rr, "I_P"] \arrow[d] & & {[0, +\infty]} \\
{[0, 1]} \arrow[urr, "I"'] & &
\end{tikzcd}
\tag{3.1}
$$

is commutative, that is, $I \circ P = I_P$, and I is continuous;

(ii) $A \parallel B \Rightarrow I_P(A \cap B) = I_P(A) + I_P(B)$, if $A, B \in \mathcal{A}$.

It is a very well-known fact that the only real function I_P which satisfies the conditions (3.1) is $I_P(A) = -\alpha \cdot \log P(A)$, where α is an arbitrary positive real constant. It is a matter of convention to choose a unit for measurement of the amount of information which makes $\alpha = 1$.

Let us now assume that we are given several experiments in the form of a system of mutually exclusive and collectively exhaustive events, $\mathcal{P} = \{A_i\}_{i=1}^{n}$, where

$$\bigcup_{i=1}^{m} A_i = \Omega, \quad \text{and} \quad A_i \cap A_j = \emptyset \quad \text{for} \quad i \neq j \quad \text{and} \quad i, j \leq m \; ^{*)} .$$

---

*) Variables i, j, k, l will always run over the set of positive natural numbers $\{1, 2, 3, \ldots\}$.

What we may well ask is the <u>average amount of information conveyed</u> by the system of experiments $\wp$.

Since we are assuming the probabilistic frame $\mathbb{A}$, there is nothing more natural than to take the average amount of information, called the <u>entropy</u> H , to be the expected value of the amount of information:

$$H_P(\wp) \;=\; \sum_{A\in\wp} P(A)\cdot I_P(A)\,, \quad \text{where} \quad I_P(A) = -\log_2 P(A)\,. \qquad (3.2)$$

The entropy measure H is usually characterized by a system of functional equations using more or less plausible ideas about the properties of H.

Let $\mathbb{P}$ be the set of all possible partitions of the basic set of elementary events Ω of the structure $\mathbb{A}$. The elements of $\mathbb{P}$ will be called for simplicity <u>experiments</u>, for $\wp \in \mathbb{P}$, where $A \in \wp$ is an event, representing a possible realization of the experiment $\wp$. Then the functional equations for H have the following form:

(i) The diagram

$$\begin{tikzcd} \mathbb{P} \arrow[r, "H_P"] \arrow[d, "{< P, P, \ldots, P >}"'] & {[0, +\infty]} \\ {[0,1]\times[0,1]\times\cdots\times[0,1]} \arrow[ur, "H"'] & \end{tikzcd} \qquad (3.3)$$

is commutative, that is, $H \circ < P, P, \ldots, P > = H_P$, and H is continuous;

(ii) $H(\{A, \bar{A}\}) = 1$ if $P(A) = P(\bar{A})$;

(iii) $H([B \mid A \cap B, \bar{A} \cap B]\mathcal{P}) = H(\mathcal{P}) + P(B) \cdot H(\{A, \bar{A}\})$ if $A \perp\!\!\!\perp B$;

here $A, B \in \mathcal{A}$ and $[B \mid A \cap B, \bar{A} \cap B]\mathcal{P}$ is the experiment which is the result of replacing B in the partition $\mathcal{P}$ by two disjoint events $A \cap B$, $\bar{A} \cap B$. It is assumed, of course, that $B \in \mathcal{P}$.

It was Fadeev [39] who showed, using Erdős' famous number-theoretic lemma about additive arithmetic functions (see Erdős [57]), that the only function H_P which satisfies the conditions (3.3) has the form (3.2).

What has been said so far is pretty standard and well known. In the sequel we shall point out a different and probably new approach. Instead of constructing functional equations and by proving the validity of the formula (3.2) and showing that they adequately mirror our ideas about the concepts of information and entropy, we propose here to approach the problem qualitatively.

Following de Finetti, Savage [6], and others, we shall assume that our probabilistic frame is a qualitative probability structure (FQP-structure) $< \Omega , \mathcal{A} , \preceq >$, where $A \preceq B$ means that the event A is not more probable that the event B $(A, B \in \mathcal{A})$. In the general case there is no need to associate the binary relation $\preceq$ with any subjectivist interpretation of probability.

The question arises whether we can introduce a binary relation $\preceq$ on the set of experiments $\mathbb{P}$ in such a way that this relation will express satisfactorily our intuitions and experiences about the notion of entropy. In other words, we would like to say under what conditions on $\preccurlyeq$ we have:

$\mathcal{P}_1 \preccurlyeq \mathcal{P}_2 \longleftrightarrow$ Experiment $\mathcal{P}_1$ does not have more entropy than the experiment $\mathcal{P}_2$.

In a way this question belongs to measurement theory (see Suppes and Zinnes [58]). When we study any property of a given family of empirical objects, or a relation among these objects, one of the basic epistemological problems is to find under what conditions the given property or relation is measurable; more specifically, what are the necessary and sufficient conditions for there to exist a real valued function on the family of empirical objects whose range is a homomorphic image of the set of empirical objects in accordance with the given property or relation?

In the case of entropy this amounts to knowing the restrictions to be imposed on $\preccurlyeq$ in order that H_P of (3.2) exists and furthermore satisfies the following homomorphism condition:

$$\mathcal{P}_1 \preccurlyeq \mathcal{P}_2 \Longleftrightarrow H_P(\mathcal{P}_1) \leq H_P(\mathcal{P}_2) \text{ , if } \mathcal{P}_1, \mathcal{P}_2 \in \mathbb{P} \tag{3.4}$$

It is a trivial matter to notice that the relation $\preccurlyeq$ has to be reflexive, transitive, connected, and antisymmetric with respect to the relation $\doteq$ (defined by $\mathcal{P}_1 \doteq \mathcal{P}_2 \Longleftrightarrow \mathcal{P}_1 \preccurlyeq \mathcal{P}_2$ & $\mathcal{P}_2 \preccurlyeq \mathcal{P}_1$, if $\mathcal{P}_1, \mathcal{P}_2 \in \mathbb{P}$). In other words, $\preccurlyeq$ has at least to be a linear ordering modulo the relation $\doteq$. But these trivial assumptions are obviously insufficient to guarantee the existence of so complicated a function as H_P .

Likewise we can introduce a binary relation $\precsim^{\circ}$ on the Boolean algebra $\mathfrak{A}$, and consider the intended interpretation

$A \precsim^{\circ} B \Leftrightarrow$ Event A does not convey more information than event B.

Again, we shall try to formulate the conditions on $\precsim^{\circ}$ which allow us to find an information function I_P (H_P) satisfying both (3.1), and the following homomorphism condition:

$$A \precsim^{\circ} B \Leftrightarrow I_P(A) \leq I_P(B) \text{ , if } A, B \in \mathfrak{A} \qquad (3.5)$$

Hence our problem is to discover some conditions which, though expressible in terms of $\precsim^{\circ}$ ($\precsim$) only, allow us to find a function I_P (H_P) satisfying (3.1), (3.5) ((3.3), (3.4)).

This approach is interesting not only theoretically but also from the point of view of applications. In social, behavioral, economic, and biological sciences there is quite often no plausible way of assigning probabilities to events. But the subject or system in question may be pretty well able to order the events according to their probabilities, informations, or entropies in a certain qualitative sense.

Of course, it is an empirical problem whether the qualitative probability, information, or entropy determined by the given subject or system then actually satisfies the required axioms. But in any case, the qualitative approach gives the measurability conditions for the analyzed probabilistic or information-theoretic property.

3.3. Basic Operations on the Set of Probabilistic Experiments

In Section 3.2 we stated that the main algebraic entity to be used in the definition of an entropy structure is the partition of the set of elementary events Ω . We decided to call partitions experiments and the set of all possible experiments over Ω has been denoted by $\mathbb{P}$. For technical reasons we shall assume sometimes that every partition contains the impossible event $\emptyset$.

We can, alternatively, analyze qualitative entropy in terms of Boolean algebras generated by experiments (partitions of the sample space). Experiments are the s ts of atoms of these Boolean algebras, and there is therefore a one-one correspondence between them. Formally we get nothing new.

If we are given two partitions $\mathcal{P}_1$, $\mathcal{P}_2$, we can define the so-called finer-than relation ($\Subset$) between them as follows:

$$\mathcal{P}_1 \Subset \mathcal{P}_2 \Leftrightarrow \forall_{A \in \mathcal{P}_1} \exists_{B \in \mathcal{P}_2} (A \subseteq B) \tag{3.6}$$

An equivalent definition would be:

$$\mathcal{P}_1 \Subset \mathcal{P}_2 \longleftrightarrow \forall_A \left(A \in \mathcal{P}_2 \Longrightarrow A = \bigcup_{i \leq k} B_i\right) \text{ for some } B_i\text{'s}$$

from $\mathcal{P}_1$, $i \leq k$.

We have in particular,

$$\ldots \Subset \{\emptyset, \bar{A}, A\bar{B}, AB\bar{C}, ABC\} \Subset \{\emptyset, \bar{A}, A\bar{B}, AB\} \Subset \{\emptyset, \bar{A}, A\} .$$

Now, given a relation on a set, it is natural to ask whether it is possible to define some kind of lattice operations induced by this relation. The answer here is positive. The first operation of interest is called the product of experiments:

$$\mathcal{P}_1 \cdot \mathcal{P}_2 = \{A \cap B : A \in \mathcal{P}_1 \ \& \ B \in \mathcal{P}_2\} \quad (\mathcal{P}_1, \mathcal{P}_2 \in \mathbb{P}) \tag{3.7}$$

or, more generally,

$$\prod_{i \leq n} \mathcal{P}_i = \{\bigcap_{i \leq n} A_i : \forall_{i \leq n} (A_i \in \mathcal{P}_i)\} .$$

$\mathcal{P}_1 \cdot \mathcal{P}_2$ is the <u>greatest</u> <u>experiment</u> which is finer than both $\mathcal{P}_1$ and $\mathcal{P}_2$; that is,

(i) $\mathcal{P}_1 \cdot \mathcal{P}_2 \subseteq \mathcal{P}_1 \ \& \ \mathcal{P}_1 \cdot \mathcal{P}_2 \subseteq \mathcal{P}_2$,

(ii) $\mathcal{P} \subseteq \mathcal{P}_1 \ \& \ \mathcal{P} \subseteq \mathcal{P}_2 \Rightarrow \mathcal{P} \subseteq \mathcal{P}_1 \cdot \mathcal{P}_2$.

Obviously $\mathcal{P}_1 \subseteq \mathcal{P}_2 \Leftrightarrow \mathcal{P}_1 \cdot \mathcal{P}_2 = \mathcal{P}_1$.

The dual operation is called the <u>sum</u> <u>of</u> <u>experiments</u> and is defined as follows:

$$\mathcal{P}_1 + \mathcal{P}_2 = \prod_{\substack{\mathcal{P}_1 \subseteq \mathcal{P} \\ \mathcal{P}_2 \subseteq \mathcal{P}}} \{\mathcal{P}\} ,$$ where $\prod$ denotes the standard generalization of the operation $\cdot$ to sets of experiments. A more concrete definition is the following:

$$\mathcal{P}_1 + \mathcal{P}_2 = \{\bigcup_{i \leq n} A_i : A_1 \not\perp A_2 \not\perp \cdots \not\perp A_n \text{ is a maximal chain of overlapping events in } \mathcal{P}_1 \cup \mathcal{P}_2\} ,$$ where $A_i \not\perp A_j \Leftrightarrow \neg A_i \perp A_j$, $i, j \leq n$.

$\mathcal{P}_1 + \mathcal{P}_2$ is the <u>smallest</u> <u>experiment</u> coarser than both $\mathcal{P}_1$ and $\mathcal{P}_2$; that is,

(i) $\mathcal{P}_1 \subseteq \mathcal{P}_1 + \mathcal{P}_2$ & $\mathcal{P}_2 \subseteq \mathcal{P}_1 + \mathcal{P}_2$,

(ii) $\mathcal{P}_1 \subseteq \mathcal{P}$ & $\mathcal{P}_2 \subseteq \mathcal{P} \Longrightarrow \mathcal{P}_1 + \mathcal{P}_2 \subseteq \mathcal{P}$.

Again it is clear that $\mathcal{P}_1 \subseteq \mathcal{P}_2 \Longleftrightarrow \mathcal{P}_1 + \mathcal{P}_2 = \mathcal{P}_2$.

The partition $\mathcal{O} = \{\emptyset, \Omega\}$ is called the <u>maximal</u> <u>experiment</u> and the partition $\mathcal{A} = \{\{\omega\} : \omega \in \Omega\} \cup \{\emptyset\}$ is called the <u>minimal</u> <u>experiment</u>. Clearly $\mathcal{A} \subseteq \mathcal{P} \subseteq \mathcal{O}$ for any $\mathcal{P} \in \mathbb{P}$. Equally straightforward are

$$\mathcal{P} \cdot \mathcal{O} = \mathcal{P} \quad \text{and} \quad \mathcal{P} + \mathcal{O} = \mathcal{O},$$
$$\mathcal{P} \cdot \mathcal{A} = \mathcal{A} \quad \text{and} \quad \mathcal{P} + \mathcal{A} = \mathcal{P}.$$

The total number of experiments e_n over a finite set Ω with n elements is given by the following recursive formula:

$$e_0 = 1 \;\&\; e_{n+1} = \sum_{i=0}^{n} \binom{n}{i} e_i .$$

The reader can easily check that the structure $< \mathbb{P}, \mathcal{O}, \mathcal{A}, +, \cdot, \subseteq >$ satisfies the lattice axioms. Unfortunately, it is not a Boolean algebra, so there is no hope of getting any useful entropy measure on it without further assumptions. The help will come from the independent relation $\perp\!\!\!\perp$ on experiments.

The structure $< \Omega, \mathbb{P}, \subseteq >$ in which the product and sum of experiments are defined will be called the <u>algebra</u> <u>of</u> <u>experiments</u>

over the set of elementary events Ω . The reader may be familiar with the following chain of isomorphisms:

$\mathbb{P} \cong$ Lattice of equivalence relations on $\Omega \cong$ Lattice of complete Boolean subalgebras of $\mathcal{U} \cong$ Lattice of subgroups of a finite group $\cong$ Lattice of subgraphs of a topological graph $\cong$ Finite geometric system of lines and pencils $\cong$ Lattice of modal operators on $\mathcal{U}$ satisfying the modal axiom system S_5.

Any one of these structures could be used as the underlying algebraic structure of the entropy measure. For example, in graph representation, the entropy measure could be viewed also as a measure of the relative complexity of graphs:

$$H_{\rho}(\mathcal{G}) = -\sum_{A \in \rho} P(A) \cdot \log_2 P(A) \ , \quad \text{where} \quad P(A) = \frac{|A|}{|V|} \ , \quad A \in \rho \ ,$$

and ρ is the partition of the set of vertices V of the graph $\mathcal{G}$. In the same way we can talk about the complexity of a group. By the complexity of a mathematical structure we mean here a function of all the elements of a (complete) set of invariants of the given structure.

3.4. Independent Experiments

DEFINITION 5 Let $\mathbb{Q} = \langle \Omega, \mathcal{U}, \prec \rangle$ be a FAQQP-structure and let $\langle \Omega, \mathbb{P}, \underline{\subset} \rangle$ be the algebra of experiments over Ω . Then we shall say that two experiments are independent, $\rho_1 \parallel \rho_2$, if and only if

$$A \in \rho_1 \ \& \ B \in \rho_2 \Rightarrow A \parallel B \ , \quad \text{for all} \quad A, B \in \mathcal{U} \ .$$

Some of the basic properties of independent experiments are stated in the following theorem.

THEOREM 13   If $<\Omega, \mathcal{A}, \precsim>$ is a FAQQP-structure modulo $\sim$ and $<\Omega, \mathbb{P}, \subseteq>$ is the algebra of experiments over Ω, then the following formulas are valid for all $\mathcal{P}, \mathcal{P}_1, \mathcal{P}_2 \in \mathbb{P}$:

(1) $\mathcal{O} \parallel \mathcal{P}$;

(2) $\mathcal{P} \parallel \mathcal{P} \Rightarrow \mathcal{P} = \mathcal{O}$;

(3) $\mathcal{P}_1 \parallel \mathcal{P}_2 \Leftrightarrow \mathcal{P}_2 \parallel \mathcal{P}_1$;

(4) $\mathcal{P}_1 \parallel \mathcal{P}_2 \ \& \ \mathcal{P}_2 \subseteq \mathcal{P}_3 \Rightarrow \mathcal{P}_1 \parallel \mathcal{P}_3$;

(5) $\mathcal{P} \parallel \mathcal{P} \Rightarrow \mathcal{P} \parallel \mathcal{P}_1$;

(6) $\mathcal{P}_1 \parallel \mathcal{P}_2 \ \& \ \mathcal{P}_2 \parallel \mathcal{P}_3 \Rightarrow (\mathcal{P}_1 \cdot \mathcal{P}_2 \parallel \mathcal{P}_3 \Leftrightarrow \mathcal{P}_1 \parallel \mathcal{P}_2 \cdot \mathcal{P}_3)$;

(7) $\mathcal{P}_1 \parallel \mathcal{P} \ \& \ \mathcal{P}_2 \parallel \mathcal{P} \Rightarrow \mathcal{P}_1 \cdot \mathcal{P}_2 \parallel \mathcal{P}$, if $A \cup B = \Omega$,

$A \in \mathcal{P}_1$, $B \in \mathcal{P}_2$;

(8) $\mathcal{P}_1 \parallel \mathcal{P}_2 \cdot \mathcal{P}_3 \ \& \ \mathcal{P}_1 \cdot \mathcal{P}_2 \parallel \mathcal{P}_3 \Rightarrow (\mathcal{P}_1 \parallel \mathcal{P}_2 \quad \mathcal{P}_2 \parallel \mathcal{P}_3)$;

(9) $\mathcal{P} \parallel \mathcal{P}_1 \ \& \ \mathcal{P} \parallel \mathcal{P}_2 \Rightarrow (\mathcal{P} \cdot \mathcal{P}_1 = \mathcal{P} \cdot \mathcal{P}_2 \Leftrightarrow \mathcal{P}_1 = \mathcal{P}_2)$;

(10) $\mathcal{P}_1 \parallel \mathcal{P}_2 \ \& \ \mathcal{P}_1 \subseteq \mathcal{P}_2 \Rightarrow \mathcal{P}_2 = \mathcal{O}$.

The proof is a simple application of Theorem 6. The assumption that $<\Omega, \mathcal{A}, \precsim>$ is a FAQQP-structure is inessential. We could as well assume any FQCP-structure or even any other structure in which the relation $\parallel$ is defined for events.

The reader will notice that the relation $\parallel$ on $\mathbb{P}$ is not unlike the disjointness relation $\perp$ on $\mathcal{A}$. In particular,

$\bar{A} = \bigcap\{B : A \cup B = \Omega \;\&\; A \perp B\}$. If we define similarly:

$$\bar{\mathcal{P}} = \bigsqcap_{\substack{\mathcal{P}\cdot\mathcal{Q} = \mathcal{A} \\ \mathcal{P} \| \mathcal{Q}}} \{\mathcal{Q}\} \quad \text{and} \quad \mathcal{P}_1 \wedge \mathcal{P}_2 = (\bar{\mathcal{P}}_1 \cdot \bar{\mathcal{P}}_2)^{-} , \quad \text{then}$$

we get a Boolean algebra of those experiments, for which $\bar{\mathcal{P}}$ exists. If $\bar{\mathcal{P}}$ exists, then it is uniquely determined, as we can easily check using Theorem 13(9). Analogously, $\mathcal{P}_1 \wedge \mathcal{P}_2$ is uniquely determined, provided that it exists.

It is unfortunate that the independence relation $\|$ on $\mathbb{P}$ generates a Boolean algebra which is only a proper subset of the lattice $\mathbb{P}$. We would hardly want to rule out those experiments which have no complements according to definition given above; for the entropy measure H_P is defined on the whole set $\mathbb{P}$. On the other hand, it is highly desirable to have on $\mathbb{P}$ a richer structure than a lattice.

In the following chapter we shall make some use of the 'partial' Boolean algebra $< \mathbb{P}, \cdot, \wedge, -, \mathcal{O}, \mathcal{A} >$ in the Representation Theorem of qualitative entropy structures.

3.5. Qualitative Entropy Structures

As already mentioned in Section 3.2, we shall develop here a qualitative theory of entropy based on qualitative probability theory. The only primitive notions used will be the qualitative entropy relation $\preccurlyeq$ and the independence relation $\|$, both relations over $\mathbb{P}$, the set of experiments.

In this chapter we shall use the following notation: If $A \in \mathcal{U}$, the experiment $\{A, \bar{A}\}$ is called a Bernoulli experiment; the variables $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2, \ldots$ will run over Bernoulli experiments. Familiar enough is the fact that each experiment $\mathcal{P} \in \mathbb{P}$ can be written as a product $\prod_{i \leq n} \mathcal{B}_i$, where the family $\{\mathcal{B}_i\}_{i \leq n}$ is so chosen that no subset of it is sufficient for the job. This representation, unfortunately, is not unique.

Experiment $\mathcal{P}$ is called (locally) equiprobable if and only if $\forall A, B \in \mathcal{U}\ (A, B \in \mathcal{P} \Longrightarrow A \sim B)$. The variables for equiprobable experiments will be $\mathcal{E}, \mathcal{E}_1, \mathcal{E}_2, \ldots$.

We call two experiments equivalent modulo $\sim$, in symbols, $\mathcal{P}_1 \sim \mathcal{P}_2$, if and only if $\mathcal{P}_2 = [B_1 | A_1]([B_2 | A_2](\ldots([B_n | A_n]\mathcal{P}_1)\ldots))$, where $\forall_{i \leq n}(A_i \sim B_i \ \&\ B_i \in \mathcal{P}_1)$ and the right-hand side of the equation is an experiment. The relation $\sim$ on $\mathbb{P}$ is clearly an equivalence relation.

We define as before:

$$\mathcal{P}_1 \dot{\sim} \mathcal{P}_2 \Longleftrightarrow (\mathcal{P}_1 \dot{\preceq} \mathcal{P}_2 \ \&\ \mathcal{P}_2 \dot{\preceq} \mathcal{P}_1),$$

$$\mathcal{P}_1 \dot{\prec} \mathcal{P}_2 \Longleftrightarrow \neg\, \mathcal{P}_2 \dot{\preceq} \mathcal{P}_1$$

$$\mathcal{O} = \{\emptyset, \Omega\},$$

$$\mathcal{A} = \{\{\omega\}: \omega \in \Omega\}.$$

Let $\mathcal{P} = \mathcal{P}_1 \odot \mathcal{P}_2 \Longleftrightarrow (\mathcal{P} = \mathcal{P}_1 \cdot \mathcal{P}_2 \ \&\ \mathcal{P}_1 \perp\!\!\!\perp \mathcal{P}_2 \ \&\ \neg\, \mathcal{P}_1, \mathcal{P}_2 \sim \mathcal{O})$; let $\mathbb{B} = \{\mathcal{P} \in \mathbb{P} : \neg\, \exists_{\mathcal{P}_1, \mathcal{P}_2}[\mathcal{P} = \mathcal{P}_1 \odot \mathcal{P}_2]\}$; and let $\mathcal{D}$ enumerate

$\mathbb{B}$, so that $\mathbb{B} = \{\mathcal{D}_i\}_{i \leq k}$. Then we define:

$\hat{\mathcal{P}} = < d_1, d_2, \ldots, d_k >$, where, if $\mathcal{P} = \mathcal{D}_i$ $(i \in \{1, 2, \ldots, k\})$, then $d_i = 1$, and $\forall_{j \neq i}$ $\& \; 1 \leq j \leq k(d_j = 0)$; otherwise $\hat{\mathcal{P}} = \hat{\mathcal{Q}}_1 + \hat{\mathcal{Q}}_2$ for some $\mathcal{Q}_1, \mathcal{Q}_2 \in \mathbb{P}$. Let $\hat{\mathcal{O}} = < 0, 0, \ldots, 0 >$ be the zero vector. In other words, $\hat{\mathcal{P}} \in \mathcal{V}(\mathbb{B})$, where $\mathbb{B}$ is the basis of the k-dimensional vector space $\mathcal{V}(\mathbb{B})$. Now we are ready for the following definition:

DEFINITION 6 Let $\mathbb{Q} = < \Omega, \mathcal{U}, \preccurlyeq >$ be a FAQQP-structure or a FQCP-structure. Then the quadruple $< \Omega, \mathbb{P}, \preccurlyeq\!\cdot, \perp\!\!\!\perp >$ is said to be a finite qualitative quasi-entropy structure (FQQE-structure) over $\mathbb{Q}$ if and only if the following conditions are satisfied for all variables running over $\mathbb{P}$:

E_0 $\mathbb{P}$ is the algebra of finite experiments over Ω ; $\perp\!\!\!\perp$ denotes the probabilistic independence relation on $\mathbb{P}$, and $\preccurlyeq\!\cdot$ is a binary relation on $\mathbb{P}$;

E_1 $\mathcal{P}_1 \subseteq \mathcal{P}_2 \Rightarrow \mathcal{P}_2 \preccurlyeq\!\cdot\, \mathcal{P}_1$;

E_2 $\mathcal{O} \preccurlyeq\!\cdot\, \mathcal{B}$, if $\mathcal{B} \dot{\sim} \mathcal{E}$;

E_3 $\mathcal{P}_1 \sim \mathcal{P}_2 \Rightarrow \mathcal{P}_1 \dot{\sim} \mathcal{P}_2$;

E_4 $\mathcal{P}_1 \preccurlyeq\!\cdot\, \mathcal{P}_2 \vee \mathcal{P}_2 \preccurlyeq\!\cdot\, \mathcal{P}_1$;

E_5 $\forall_{i < n}(\mathcal{P}_i \preccurlyeq\!\cdot\, \mathcal{Q}_i) \rightarrow \mathcal{P}_n \preccurlyeq\!\cdot\, \mathcal{Q}_n$, if $\sum_{i \leq n} \hat{\mathcal{P}}_i = \sum_{i \leq n} \hat{\mathcal{Q}}_i$.

Remarks:

(1) In axiom E_5, the formula concerning characteristic functions can easily be translated into a system of identities among experiments.

(ii) There is no doubt that the axioms E_0 - E_5 are consistent and independent. The crucial axioms are E_4 and E_5. Axioms E_1 and E_2 give the so-called normalization conditions, whereas E_3 forces us to consider equiprobable classes of events, rather than events themselves.

(iii) The definition of an <u>infinite</u> <u>qualitative</u> <u>quasi</u>-<u>entropy</u> <u>structure</u> for purposes of representation by an entropy measure on $\mathbb{P}$ does not cause any fundamental difficulties. The FAQQP-structure or the FQCP-structure must of course be replaced by an infinite one; otherwise, we proceed as in the finite case. In fact, the axioms are prodigiously complicated and far less intuitive than those given above. This case will be omitted here.

(iv) In axiom E_2, we must assume the existence of an equiprobable experiment $\mathcal{E}$, for we need this axiom to show that the entropy measure H is strictly positive for at least one element from $\mathbb{P}$. An alternative axiom might be $\mathcal{O} \prec\cdot \mathcal{A}$, but this would rule out some elementary algebras $\mathbb{P}$.

(v) Note that the (<u>global</u>) <u>entropy</u> <u>relation</u> $\prec\cdot$ depends on two factors: on the underlying <u>algebra</u> <u>of</u> <u>experiments</u> and the <u>independence</u> <u>relation</u> $\parallel$ defined on this algebra (this relation is hidden in axiom E_5). We do not give here the link between the FQQE-structure (macro-structure) and the FAQQP-structure (micro-structure).

The following easy theorem displays the content of the above definition:

THEOREM 14 <u>Let</u> $< \Omega, \mathbb{P}, \preccurlyeq\cdot, \parallel >$ <u>be</u> <u>a</u> FQQE-structure <u>or</u> <u>a</u> FQCP-structure. <u>Then</u> <u>for</u> <u>all</u> <u>variables</u> <u>running</u> <u>over</u> $\mathbb{P}$ <u>and</u> $A, B, C \in \mathcal{A}$:

(1) $\mathcal{P} \preccurlyeq\cdot \mathcal{P}$;

(2) $\{\emptyset\} \cup \mathcal{P} \dot{\sim} \mathcal{P}$;

(3) $\mathcal{P}_1 \preccurlyeq\cdot \mathcal{P}_2 \ \& \ \mathcal{P}_2 \preccurlyeq\cdot \mathcal{P}_3 \Rightarrow \mathcal{P}_1 \preccurlyeq\cdot \mathcal{P}_3$;

(4) $\mathcal{P}_1 \preccurlyeq\cdot \mathcal{P}_2 \ \& \ \mathcal{P}_2 \preccurlyeq\cdot \mathcal{P}_1 \Rightarrow \mathcal{P}_1 \dot{\sim} \mathcal{P}_2$;

(5) $\sim$ <u>is</u> <u>an</u> <u>equivalence</u> <u>relation</u>;

(6) $\mathcal{O} \preccurlyeq\cdot \mathcal{P}$;

(7) $\mathcal{O} \preccurlyeq\cdot \{A, \overline{A}\} \preccurlyeq\cdot \{A, \overline{A}B, \overline{AB}\} \preccurlyeq\cdot \{A, \overline{A}B, \overline{AB}C, \overline{ABC}\} \preccurlyeq\cdot \ldots \preccurlyeq\cdot \mathcal{A}$;

(8) $\mathcal{P}_1 \preccurlyeq\cdot \mathcal{P}_2 \Leftrightarrow \mathcal{P}_1 \cdot \mathcal{P} \preccurlyeq\cdot \mathcal{P}_2 \cdot \mathcal{P}$, if $\mathcal{P} \parallel \mathcal{P}_1, \mathcal{P}_2$;

(9) $\mathcal{P}_1 \preccurlyeq\cdot \mathcal{Q}_1 \ \& \ \mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_2 \rightarrow \mathcal{P}_1 \cdot \mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_1 \cdot \mathcal{Q}_2$, <u>if</u> $\mathcal{P}_1 \parallel \mathcal{P}_2 \ \& \ \mathcal{Q}_1 \parallel \mathcal{Q}_2$;

(10) $\mathcal{P}_1 \dot{\sim} \mathcal{P}_2 \ \& \ \mathcal{P}_2 \cdot \mathcal{P}_3 \preccurlyeq\cdot \mathcal{P}_4 \rightarrow \mathcal{P}_1 \cdot \mathcal{P}_3 \preccurlyeq\cdot \mathcal{P}_4$, <u>if</u> $\mathcal{P}_3 \parallel \mathcal{P}_1, \mathcal{P}_2$;

(11) $\forall_{i < n}(\mathcal{P}_i \preccurlyeq\cdot \mathcal{Q}_i) \rightarrow \mathcal{Q}_n \preccurlyeq\cdot \mathcal{P}_n$, <u>if</u>

$$\prod_{i \leq n} \mathcal{P}_i \dot{\sim} \prod_{i \leq n} \mathcal{Q}_i \ \& \ \mathcal{P}_i \underset{i \leq n}{\parallel} \ \& \ \mathcal{Q}_i \underset{i \leq n}{\parallel} ;$$

(12) $\forall_{i \leq n}(A_i \sim B_i) \Rightarrow \{A_i\}_{i \leq n} \dot{\sim} \{B_i\}_{i \leq n}$, if $\{A_i\}_{i \leq n}$, $\{B_i\}_{i \leq n} \in \mathbb{P}$.

The empirical content of Theorem 14 should be clear.

THEOREM 15 (Representation Theorem) Let $< \Omega, \mathbb{P}, \dot{\preceq}, \| >$ be a structure, where Ω is a nonempty finite set; $\mathbb{P}$ is the set of partitions of Ω ; $\|$ is the independence relation on $\mathbb{P}$ in the sense of Definition 5; and $\dot{\preceq}$ is a binary relation on $\mathbb{P}$.

Then $< \Omega, \mathbb{P}, \dot{\preceq}, \| >$ is a FQQE-structure if and only if there exists a quasi-entropy function $H : \mathbb{P} \longrightarrow Re$ satisfying the following conditions for all $\mathcal{P}_1, \mathcal{P}_2 \in \mathbb{P}$:

(i) $\mathcal{P}_1 \dot{\preceq} \mathcal{P}_2 \Longleftrightarrow H(\mathcal{P}_1) \leq H(\mathcal{P}_2)$;

(ii) $\mathcal{P}_1 \| \mathcal{P}_2 \Longrightarrow H(\mathcal{P}_1 \cdot \mathcal{P}_2) = H(\mathcal{P}_1) + H(\mathcal{P}_2)$;

(iii) $\mathcal{P}_1 \Subset \mathcal{P}_2 \longrightarrow H(\mathcal{P}_2) \leq H(\mathcal{P}_1)$;

(iv) $H(\mathcal{O}) = 0$;

(v) $H(\mathcal{B}) = 1$, if $\mathcal{B} \dot{\sim} \mathcal{E}$.

Proof: There is no question of the conditions' not being necessary, and we prove here only their sufficiency.

Let $< \Omega, \mathbb{P}, \dot{\preceq}, \| >$ be a FQQE-structure over $\mathbb{Q}$. Let $\mathcal{V}(\mathbb{B})$ be the k-dimensional vector space, described just before Definition 6. We can obviously make $\mathbb{P}$ a finite subset of $\mathcal{V}(\mathbb{B})$

by assigning to each $\mathcal{P} \in \mathbb{P}$ a vector $\hat{\mathcal{P}}$, where $(\mathcal{P}_1 \odot \mathcal{P}_2)^\wedge = \hat{\mathcal{P}}_1 + \hat{\mathcal{P}}_2$. In a similar way $\precsim$ can be represented on $\mathcal{V}(\mathbb{B})$. Having done this, we are ready to use Corollary 5 taking advantage of E_4, E_5, and E_3 to switch to quotient structures. The corollary answers us that there is a linear functional $\psi : \mathcal{V}(\mathbb{B}) \longrightarrow Re$, and thus another functional $\varphi : \mathbb{P} \longrightarrow Re$, such that the conditions (i), (ii), and (iv) of Theorem 15 are satisfied by φ. E_1 forces φ to be non-negative on $\mathbb{P}$, and also to satisfy (iii).

Finally, E_2 gives $\varphi(\{A, \bar{A}\}) > 0$, if $E \sim \bar{E}$. Hence, by putting

$$H(\mathcal{P}) = \frac{\varphi(\mathcal{P})}{\varphi(\{A, \bar{A}\})} ,$$

we get the desired quasi-entropy function. Q. E. D.

Condition (iii) in Theorem 15 expresses the most important property of the entropy measure, namely, its additivity. Unfortunately, this property is much weaker than (iii) in (3.3), Section 3.2. It is trivial to show that there are many functions besides (3.2) which satisfy the above conditions. This lack of specificity explains the 'quasi-' prefix.

It is well known that the conditions (i) - (v) in Theorem 15 together with the condition

$$H(\mathcal{P}) = \sum_{A \in \mathcal{P}} f \circ P(A) , \quad f : [0, 1] \longrightarrow Re , \quad f(\tfrac{1}{2}) = \tfrac{1}{2} , \qquad (3.8)$$

for some f continuous, are enough to specify an entropy measure H_P in the form (3.2).

In order to guarantee the existence of a continuous function f, satisfying (3.8), we have further to restrict $\preccurlyeq^{\cdot}$, and to add more 'interacting' conditions between $\preccurlyeq^{\cdot}$ and $\preccurlyeq$.

The following necessary conditions are obvious candidates:

(1) $A \preccurlyeq B \Leftrightarrow \{A, \bar{A}\} \preccurlyeq^{\cdot} \{B, \bar{B}\}$, if $A, B \preccurlyeq E \sim \bar{E}$;

(2) $A \preccurlyeq B \Leftrightarrow \{B, \bar{B}\} \preccurlyeq^{\cdot} \{A, \bar{A}\}$, if $\bar{E} \sim E \preccurlyeq A, B$;

(3) $A \preccurlyeq C \Leftrightarrow [B \mid AB, \bar{A}B]\mathscr{P} \preccurlyeq^{\cdot} [B \mid CB, \bar{C}B]\mathscr{P}$*) , if $\emptyset \prec B \in \mathscr{P}$, $B, C \preccurlyeq E \sim \bar{E}$, $\emptyset \prec B$, $C \prec \Omega$, $A \parallel B, C$;

(4) $B \preccurlyeq C \Leftrightarrow [B \mid AB, \bar{A}B]\mathscr{P} \preccurlyeq^{\cdot} [C \mid AC, \bar{A}C]\mathscr{P}$, if $B, C \in \mathscr{P}$, $A \parallel B, C$, $\emptyset \prec A \prec \Omega$;

(5) $\mathscr{P} \preccurlyeq^{\cdot} \mathscr{E}$, if $|\mathscr{P}| = |\mathscr{E}|$.

It would be incredible if these conditions were sufficient. At least three axioms or axiom schemas similar to E_5 are needed to guarantee the existence of the sum, multiplication, and logarithm functions in (3.2). Over and above that we need the qualitative probability axioms, which we can assume to be given, of course.

Given these axioms, our representation theorem would also guarantee the existence of a probability measure P such that in addition to (i) - (v) in Theorem 15 we would have:

(vi) $H(\mathscr{P}) + H(\{A, \bar{A}\}) \cdot P(B) = H([B \mid AB, \bar{A}B]\mathscr{P})$, if $B \in \mathscr{P}$, and $A \parallel B$, $\emptyset \prec A \prec \Omega$;

---

*) See the notation in (3.3), Section 3.2.

(vii) $A \prec B \Leftrightarrow P(A) \leq P(B)$;

(viii) $A \parallel B \Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$;

(ix) $\mathcal{P}_1 \parallel \mathcal{P}_2 \Leftrightarrow \forall A, B(A \in \mathcal{P}_1 \ \& \ B \in \mathcal{P}_2 \Rightarrow A \parallel B)$.

It seems to be an open problem to specify the relationship between the <u>macro</u>- and <u>micro-structures</u> under the given very restrictive <u>finite</u> <u>conditions</u>. On the other hand, in the next section we shall show how easy it is to give 'representations' when we have more topological properties available.

FQQE-structures characterize the macro-properties of the entropy from qualitative point of view. The reader may have noticed the following striking formal similarity between the conditional entropy measure and the (absolute) probability measure:

(1) $\mathcal{P}_1 \parallel \mathcal{P}_2 \Rightarrow H(\mathcal{P}_1/\mathcal{P}_2) = H(\mathcal{P}_1)$,

$A \perp B \Rightarrow P(A - B) = P(A)$,

(2) $H(\mathcal{P}_1 \cdot \mathcal{P}_2) = H(\mathcal{P}_2) + H(\mathcal{P}_1/\mathcal{P}_2)$,

$P(A \cup B) = P(B) + P(A - B)$,

(3) $0 \leq H(\mathcal{P}_1) \leq H(\mathcal{P}_1 \cdot \mathcal{P}_2) \leq H(\mathcal{P}_1) + H(\mathcal{P}_2)$,

$0 \leq P(A) \leq P(A \cup B) \leq P(A) + P(B)$.

This rather primitive one-one correspondence between

$\mathcal{P}_1 \parallel \mathcal{P}_2$, $\mathcal{P}_1 \cdot \mathcal{P}_2$, $\mathcal{P}_1/\mathcal{P}_2$ and $A \perp B$, $A \cup B$, $A - B$

contains certainly some heuristic anticipation of a deeper relationship between the macro- and micro-structures: $< \Omega, \mathbb{P}, H >$ and $< \Omega, \mathfrak{A}, P >$. One can see also why the lattice operation $+$

in $\mathbb{P}$ has so little use in entropy theory. The more interesting operation on $\mathbb{P}$ would be the composition of two experiments, $\mathcal{P}_1 \wedge \mathcal{P}_2$, defined in Section 3.3. The only problem here is that $< \mathbb{P}, \mathcal{O}, \mathcal{U}, \cdot, \wedge >$ cannot be embedded into a Boolean algebra.

We shall now turn to the problem of conditional entropy. Another interesting similarity between the conditional entropy and (conditional) probability is the following:

(1) $H(\mathcal{P}_1/\mathcal{P}_2) = H(\mathcal{P}_1 \cdot \mathcal{P}_2) - H(\mathcal{P}_2)$,

$P(A/B) = P(AB)/P(B),\ P(B) > 0$,

(2) $\mathcal{P}_1 \| \mathcal{P}_2 \Longleftrightarrow H(\mathcal{P}_1/\mathcal{P}_2) = H(\mathcal{P}_1/\mathcal{O})$,

$A \| B \Longleftrightarrow P(A/B) = P(A/\Omega)$, if $P(B) > 0$,

(3) $H(\mathcal{P}_1/\mathcal{P}_2 \cdot \mathcal{P}_3) = H(\mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P}_3) - H(\mathcal{P}_2/\mathcal{P}_3)$,

$P(A/BC) = P(AB/C) \,/\, P(B/C)$, if $P(BC) \cdot P(C) > 0$.

We shall consider these similarities as a heuristic guide to further developments of entropy structures. One can consider the entity $\mathcal{P}_1/\mathcal{P}_2$ to be a partition (experiment) in $\mathfrak{A}/\mathfrak{A}[\mathcal{P}_2]$. Then $\mathcal{P}_1/\mathcal{P}_2$ is the set of experiments indistinguishable from $\mathcal{P}_1$, given $\mathcal{P}_2$.

As in the case of probability structures (see Section 2.4, Definition 2) we shall study a kind of composition of entropy structures. In particular, given the algebra of experiments $< \Omega, \mathbb{P}, \subseteq >$, we shall study a binary relation $\precsim$ on $\mathbb{P} \times \mathbb{P}$

and a special representation function $\psi : \mathbb{P} \longrightarrow \mathrm{Re}$, which, among other things, satisfies

$$< \mathcal{P}_1 , \mathcal{P}_2 > \precsim < \mathcal{Q}_1 , \mathcal{Q}_2 > \leftrightarrow \psi(\mathcal{P}_1) + \psi(\mathcal{P}_2) \leq \psi(\mathcal{Q}_1) + \psi(\mathcal{Q}_2)$$

for all $\mathcal{P}_1, \mathcal{P}_2, \mathcal{Q}_1, \mathcal{Q}_2 \in \mathbb{P}$.

There are several important partial interpretations of this relation: First of all, the <u>qualitative</u> <u>conditional</u> <u>quasi-entropy</u> <u>relation</u> hopefully can be defined as

$$\mathcal{P}_1 / \mathcal{P}_2 \precsim\cdot \mathcal{Q}_1 / \mathcal{Q}_2 \leftrightarrow < \mathcal{P}_1 \cdot \mathcal{P}_2, \mathcal{Q}_2 > \precsim < \mathcal{Q}_1 \cdot \mathcal{Q}_2, \mathcal{P}_2 > .$$

Naturally, we can put

$$\mathcal{P}_1 \precsim\cdot \mathcal{P}_2 \Leftrightarrow < \mathcal{P}_1, \mathcal{O} > \precsim < \mathcal{P}_2, \mathcal{O} >$$

and then the <u>probabilistic</u> <u>independence</u> <u>relation</u> $\parallel$ on experiments is given by

$$\mathcal{P}_1 \parallel \mathcal{P}_2 \Leftrightarrow < \mathcal{P}_1, \mathcal{P}_2 > \precsim < \mathcal{P}_1 \cdot \mathcal{P}_2, \mathcal{O} > .$$

It is clear that we could also talk about <u>positive</u> and <u>negative</u> <u>dependence</u> <u>notions</u> similar to those introduced for probabilities. The structure $< \mathbb{P} \times \mathbb{P} , \precsim >$ also has independent importance in algebraic measurement theory, where the atomic formula $< \mathcal{P}_1, \mathcal{P}_2 > \precsim < \mathcal{Q}_1, \mathcal{Q}_2 >$ may be interpreted as a comparison of two <u>empirical</u> <u>compositions</u> of certain physical entities, which is representable by an inequality between the sum of magnitudes of a linear physical quantity. In this paper we shall be interested only in the entropy-interpretation.

DEFINITION 7 Let $\mathfrak{Q} = \langle \Omega, \mathcal{A}, \precsim \rangle$ be a FAQQP-structure. Then the quadruple $\langle \Omega, \mathbb{P}, \precsim^{\bullet}, \| \rangle$ is said to be a finite qualitative quasi-entropy difference structure (FQQED-structure) over $\mathfrak{Q}$ if and only if the following conditions are satisfied for all variables running over $\mathbb{P}$:

D_0 $\mathbb{P}$ is the algebra of finite experiments over Ω; $\|$ is the probabilistic independence relation on $\mathbb{P}$ and $\precsim$ is a relation on $\mathbb{P} \times \mathbb{P}$;

D_1 $\mathcal{P}_1 \subseteq \mathcal{P}_2 \Rightarrow \langle \mathcal{P}_2, \mathcal{P} \rangle \precsim \langle \mathcal{P}_1, \mathcal{P} \rangle$;

D_2 $\langle \mathcal{O}, \mathcal{P} \rangle \precsim \langle \mathcal{B}, \mathcal{P} \rangle$, if $\mathcal{B} \simeq \mathcal{E}$;

D_3 $\langle \mathcal{P}_1, \mathcal{P}_2 \rangle \precsim \langle \mathcal{Q}_1, \mathcal{Q}_2 \rangle \vee \langle \mathcal{Q}_1, \mathcal{Q}_2 \rangle \precsim \langle \mathcal{P}_1, \mathcal{P}_2 \rangle$;

D_4 $\langle \mathcal{P}_1, \mathcal{P}_2 \rangle \precsim \langle \mathcal{Q}_1, \mathcal{Q}_2 \rangle \Rightarrow \langle \mathcal{Q}_2, \mathcal{Q}_1 \rangle \precsim \langle \mathcal{P}_2, \mathcal{P}_1 \rangle$;

D_5 $\forall_{i<n}(\langle \mathcal{P}_i, \mathcal{Q}_i \rangle \precsim \langle \mathcal{R}_i, \mathcal{S}_i \rangle) \Rightarrow \langle \mathcal{S}_n, \mathcal{R}_n \rangle \precsim \langle \mathcal{P}_n, \mathcal{Q}_n \rangle$,

if $\sum_{i \leq n} \hat{\mathcal{P}}_i = \sum_{i \leq n} \hat{\mathcal{R}}_i$ & $\sum_{i \leq n} \hat{\mathcal{Q}}_i = \sum_{i \leq n} \hat{\mathcal{S}}_i$,

where $\hat{\mathcal{P}}_i$, $\hat{\mathcal{R}}_i$, $\hat{\mathcal{Q}}_i$, $\hat{\mathcal{S}}_i$ for $i = 1, 2, \ldots, n$ have the same meaning as in Definition 6.

The remarks to Definition 6 are relevant also to Definition 7. The content of the definition should be clear; therefore we proceed to Theorem 16.

THEOREM 16 (Representation Theorem) Let $< \Omega, \mathbb{P}, \preccurlyeq, \perp\!\!\!\perp >$ be a structure, where Ω is a nonempty finite set; $\mathbb{P}$ is the set of partitions of Ω; $\perp\!\!\!\perp$ is the independence relation on in the sense of the Definition 5; and $\preccurlyeq$ is a relation on $\mathbb{P} \times \mathbb{P}$.

Then $< \Omega, \mathbb{P}, \preccurlyeq, \perp\!\!\!\perp >$ is a FQQED-structure if and only if there exists a quasi-entropy function $H : \mathbb{P} \longrightarrow Re$ satisfying the following conditions for all $\mathcal{P}_1, \mathcal{P}_2, \mathcal{Q}_1, \mathcal{Q}_2 \in \mathbb{P}$:

(i) $< \mathcal{P}_1, \mathcal{P}_2 > \preccurlyeq < \mathcal{Q}_1, \mathcal{Q}_2 > \Leftrightarrow H(\mathcal{P}_1) - H(\mathcal{P}_2) \leq H(\mathcal{Q}_1) - H(\mathcal{Q}_2)$;

(ii) H satisfies conditions (ii) - (v) of Theorem 15.

Proof: The necessity is obvious. For sufficiency, let $< \Omega, \mathbb{P}, \preccurlyeq, \perp\!\!\!\perp >$ be a FQQED-structure over $\mathbb{Q}$. Let $\mathcal{V}(\mathbb{B})$ be the k-dimensional vector space, described in the proof of Theorem 15. We can transform $\mathbb{P} \times \mathbb{P}$ into a finite subset of the (external) direct sum $\mathcal{V}(\mathbb{B}) \oplus \mathcal{V}(\mathbb{B})$ by assigning to each pair $< \mathcal{P}, \mathcal{Q} >$ a vector $\hat{\mathcal{P}} \oplus \hat{\mathcal{Q}} \in \mathcal{V}(\mathbb{B}) \oplus \mathcal{V}(\mathbb{B})$. We then proceed almost exactly as does Scott (D. Scott [11], Theorem 3.2, p. 245), so that the axioms D_3, D_4, D_5 are justified. As in Theorem 15, the normalization conditions D_1, D_2 will allow us to construct a function H (which exists on the basis of $D_3 - D_5$) with the desired properties (i) and (ii) in Theorem 16. Q. E. D.

Now if we put

$$\mathcal{P}_1/\mathcal{P}_2 \dot{\preccurlyeq} \mathcal{Q}_1/\mathcal{Q}_2 \Leftrightarrow < \mathcal{P}_1 \cdot \mathcal{P}_2, \mathcal{P}_2 > \preccurlyeq < \mathcal{Q}_1 \cdot \mathcal{Q}_2, \mathcal{Q}_2 > \qquad (3.9)$$

we can easily prove the following theorem with the help of Definition 7:

THEOREM 17 *Let* $< \Omega, \mathbb{P}, \preccurlyeq, \parallel >$ *be a* FQQED-structure over *a* FQCP-structure. *Then the following formulas hold, when all variables run over* $\mathbb{P}$:

(1) $\dot{\sim}$ *is an equivalence relation;*

(2) $\mathcal{O}/\mathcal{P} \dot{\preccurlyeq} \mathcal{P}_1/\mathcal{P}_2$;

(3) $\mathcal{O}/\mathcal{B} \dot{\preccurlyeq} \mathcal{B}/\mathcal{B}$, *if* $\mathcal{B} \dot{\sim} \mathcal{E}$;

(4) $\mathcal{P}_1/\mathcal{P}_2 \dot{\sim} \mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P}_2$;

(5) $\mathcal{P}/\mathcal{P} \dot{\sim} \mathcal{Q}/\mathcal{Q}$;

(6) $\mathcal{O}/\mathcal{P}_1 \dot{\preccurlyeq} \mathcal{P}_2/\mathcal{P}_2$;

(7) $\mathcal{P}_1 \dot{\preccurlyeq} \mathcal{P}_2 \Leftrightarrow \mathcal{P}_1/\mathcal{P}_2 \dot{\preccurlyeq} \mathcal{P}_2/\mathcal{P}_1$;

(8) $\mathcal{P}_1/\mathcal{P}_2 \cdot \mathcal{P}_3 \dot{\preccurlyeq} \mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P}_3$;

(9) $\mathcal{P}_1/\mathcal{P}_2 \cdot \mathcal{P}_3 \dot{\preccurlyeq} \mathcal{P}_1/\mathcal{P}_2$;

(10) $\mathcal{P}_1/\mathcal{P}_2 \dot{\preccurlyeq} \mathcal{P}_1$;

(11) $\mathcal{P}_1 \subseteq \mathcal{P}_2 \Rightarrow \mathcal{P}_2/\mathcal{P} \dot{\preccurlyeq} \mathcal{P}_1/\mathcal{P}$;

(12) $\mathcal{P} \subseteq \mathcal{P}_1 \Rightarrow \mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P} \dot{\sim} \mathcal{P}_2/\mathcal{P}$;

(13) $\mathcal{P}_1 \subseteq \mathcal{P}_2 \Rightarrow \mathcal{P}_2/\mathcal{P}_1 \dot{\sim} \mathcal{O}$;

(14) $\mathcal{P}_1/\mathcal{P}_2 \dot{\sim} \mathcal{P}_1 \Leftrightarrow \mathcal{P}_1 \parallel \mathcal{P}_2$;

(15) $\mathcal{P}_1/\mathcal{P} \preccurlyeq\cdot \mathcal{P}_2/\mathcal{P} \Leftrightarrow \mathcal{P}_1 \cdot \mathcal{Q}/\mathcal{P} \preccurlyeq\cdot \mathcal{P}_2 \cdot \mathcal{Q}/\mathcal{P}$, *if*

$\mathcal{P}_1, \mathcal{P}_2 \parallel \mathcal{Q}$;

(16) $\mathcal{P}_1 \cdot \mathcal{P}_2 \dot{\sim} \mathcal{Q}_1 \cdot \mathcal{Q}_2 \Rightarrow (\mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_2 \Leftrightarrow \mathcal{Q}_1/\mathcal{Q}_2 \preccurlyeq\cdot \mathcal{P}_1/\mathcal{P}_2)$;

(17) $(\mathcal{P}_1/\mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_1/\mathcal{Q}_2 \;\&\; \mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_2) \Rightarrow \mathcal{P}_1 \cdot \mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_1 \cdot \mathcal{Q}_2$;

(18) $(\mathcal{P}_1 \cdot \mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_1 \cdot \mathcal{Q}_2 \;\&\; \mathcal{Q}_2 \preccurlyeq\cdot \mathcal{P}_2) \Rightarrow \mathcal{P}_1/\mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_1/\mathcal{Q}_2$;

(19) $(\mathcal{P}_1/\mathcal{P}_2 \cdot \mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_1/\mathcal{Q}_2 \cdot \mathcal{Q}_3 \;\&\; \mathcal{P}_2/\mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_2/\mathcal{Q}_3) \Longrightarrow$

$\mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_1 \cdot \mathcal{Q}_2/\mathcal{Q}_3$;

(20) $\mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P}_3 \dot{\sim} \mathcal{Q}_1 \cdot \mathcal{Q}_2/\mathcal{Q}_3 \longrightarrow$

$(\mathcal{P}_1/\mathcal{P}_2 \cdot \mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_1/\mathcal{Q}_2 \cdot \mathcal{Q}_3 \Rightarrow \mathcal{Q}_2/\mathcal{Q}_3 \preccurlyeq\cdot \mathcal{P}_2/\mathcal{P}_3)$;

(21) $(\mathcal{P}_1/\mathcal{P}_2 \cdot \mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_2/\mathcal{Q}_3 \;\&\; \mathcal{P}_2/\mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_1/\mathcal{Q}_2 \cdot \mathcal{Q}_3) \longrightarrow$

$\mathcal{P}_1 \cdot \mathcal{P}_2/\mathcal{P}_3 \preccurlyeq\cdot \mathcal{Q}_1 \cdot \mathcal{Q}_2/\mathcal{Q}_3$;

(22) $(\mathcal{P}_1/\mathcal{P}_2 \preccurlyeq\cdot \mathcal{Q}_1/\mathcal{Q}_2 \;\&\; \mathcal{Q}_1/\mathcal{Q}_2 \preccurlyeq\cdot \mathcal{R}_1/\mathcal{R}_2) \Rightarrow \mathcal{P}_1/\mathcal{P}_2 \preccurlyeq\cdot \mathcal{R}_1/\mathcal{R}_2$;

(23) $\prod_{i=1}^{n} \mathcal{P}_i \dot{\sim} \prod_{i=1}^{n} \mathcal{Q}_i \;\&\; \forall_{1 \le i \le n}\left(\mathcal{P}_i \Big/ \prod_{j=0}^{i-1} \mathcal{P}_j \preccurlyeq\cdot \mathcal{Q}_i \Big/ \prod_{j=0}^{i-1} \mathcal{Q}_j\right) \longrightarrow$

$\mathcal{Q}_n \Big/ \prod_{j=0}^{n-1} \mathcal{Q}_j \preccurlyeq\cdot \mathcal{P}_n \Big/ \prod_{j=0}^{n-1} \mathcal{P}_j$ <u>where</u> $\mathcal{P}_0 \dot{\sim} \mathcal{O}$.

No more than with Definition 6 can we hope to show that

$$H(\mathcal{P}_1/\mathcal{P}_2) = -\sum_{A\in \mathcal{P}_1, B\in \mathcal{P}_2} P(AB) \cdot \log_2 P(A/B) \tag{3.10}$$

without giving some further axioms to link $\preccurlyeq\cdot$ with the probability relation $\preccurlyeq$ on $\mathcal{U}$.

It was Khinchin [40] who showed that the conditions

(a) $H(\mathcal{P}_1 \cdot \mathcal{P}_2) - H(\mathcal{P}_2) = H(\mathcal{P}_1 / \mathcal{P}_2)$;

(b) $H(\mathcal{P}) \leq H(\mathcal{E})$, if $|\mathcal{P}| = |\mathcal{E}|$;

(c) $H(\mathcal{P} \cup \{\emptyset\}) = H(\mathcal{P})$;

imply the identity

$$H(\mathcal{P}) = - \sum_{A \in \mathcal{P}} P(A) \cdot \log_2 P(A) ,$$

and therefore also the identity (3.10). In our case (a) is true by definition, and (b) and (c) become valid by adding the following two axioms:

D_6 $\langle \mathcal{P}, \mathcal{Q} \rangle \precsim \langle \mathcal{E}, \mathcal{Q} \rangle$, if $|\mathcal{P}| = |\mathcal{E}|$;

D_7 $\langle \mathcal{P} \cup \{\emptyset\}, \mathcal{Q} \rangle \sim \langle \mathcal{P}, \mathcal{Q} \rangle$.

Naturally $\mathcal{E}$ must exist, otherwise the axiom D_6 would be vacuously true. Given that, D_0 - D_7 imply the conditions (a), (b), (c) for <u>finite qualitative conditional entropy relations</u>.

3.6. <u>Qualitative Information Structures</u>

The reader may be somewhat disappointed after reading the previous section by the very general and rather weak nature of the results on entropy structures. It should be emphasized again, however, that we cannot expect simple results about fairly complicated continuous functions in terms of relations on <u>finite</u> domains.

In this section, unlike the earlier ones, we shall work with infinite Boolean algebras; as we shall see, the results will be somewhat stronger. We are able to give a definition of information measure without any recourse to probabilistic notions.

The structure to be studied here is a Boolean algebra $\mathcal{Ol}$ enriched by two binary relations $\|$ and $\precsim^{\circ}$; the relation $\|$ can be interpreted as follows:

$A \| B \Leftrightarrow$ Event A is independent of event B $(A, B \in \mathcal{Ol})$,

and the $\precsim^{\circ}$ is interpreted as before:

$A \precsim^{\circ} B \Leftrightarrow$ Event A does not have more information than event B $(A, B \in \mathcal{Ol})$.

The novelty here is that we give axioms for $\|$, $\precsim^{\circ}$, and $\mathcal{Ol}$ which, without recourse to probability theory, ensure the existence of an information measure in the standard sense.

The need for a formalization of a notion of qualitative independence to match the standard probabilistic notion has been felt for a long time, but the author is not aware of any serious attempts to solve this problem. In this section we shall try to work out such a formalization. First, perhaps, we should turn to the definition:

DEFINITION 8 Let Ω be a nonempty set, $\mathcal{Ol}$ a nonempty family of subsets of Ω such that it is a Boolean algebra, and $\|$ and $\precsim^{\circ}$ binary relations on $\mathcal{Ol}$.

Then the quadruple $\langle \Omega, \mathscr{U}, \prec^{\circ}, \parallel \rangle$ is called a qualitative information structure (QI-structure) if and only if the following conditions are satisfied when all variables run over $\mathscr{U}$:

I_1 $\quad \emptyset \parallel A$;

I_2 $\quad A \parallel B \Rightarrow B \parallel A$;

I_3 $\quad A \parallel B \Rightarrow \overline{B} \parallel A$;

I_4 $\quad A \parallel B \;\&\; A \parallel C \Rightarrow A \parallel B \cup C$, if $B \perp C$;

I_5 $\quad \Omega \succ^{\circ} \emptyset$;

I_6 $\quad A \succcurlyeq^{\circ} \emptyset$;

I_5 $\quad A \preccurlyeq^{\circ} B \vee B \preccurlyeq^{\circ} A$;

I_8 $\quad A \preccurlyeq^{\circ} B \;\&\; B \preccurlyeq^{\circ} C \Rightarrow A \preccurlyeq^{\circ} C$;

I_9 $\quad A \parallel B \;\&\; A \perp B \Rightarrow (A \overset{\circ}{\sim} \emptyset \vee B \overset{\circ}{\sim} \emptyset)$;

I_{10} $\quad A \preccurlyeq^{\circ} B \Leftrightarrow A \cup C \preccurlyeq^{\circ} B \cup C$, if $C \perp A, B$;

I_{11} $\quad A \preccurlyeq^{\circ} B \Leftrightarrow A \cap C \preccurlyeq^{\circ} B \cap C$, if $C \parallel A, B \;\&\; C \succ^{\circ} \emptyset$;

I_{12} $\quad A \preccurlyeq^{\circ} B \;\&\; C \preccurlyeq D \Rightarrow A \cup C \preccurlyeq^{\circ} B \cup D$, if $B \perp D$;

I_{13} $\quad A \preccurlyeq^{\circ} B \;\&\; C \preccurlyeq D \Rightarrow A \cap C \preccurlyeq^{\circ} B \cap D$, if $A \parallel C \;\&\; B \parallel D$;

I_{14} If $A_i \perp A_j$ for $i \neq j \;\&\; i, j \leq n$, then

$$\forall_B \exists_{A_{n+1}} \forall_{i \leq n} (A_i \perp A_{n+1} \;\&\; B \overset{\circ}{\sim} A_{n+1}) ;$$

I_{15} If $A_i \parallel A_j$ for $i \neq j \;\&\; i, j \leq n$, then

$$\forall_B \exists_{A_{n+1}} \forall_{i \leq n} (A_1 \parallel A_{n+1} \;\&\; B \overset{\circ}{\sim} A_{n+1}) .$$

Remarks:

(i) All axioms but the last two, which force $\mathfrak{A}$ to be infinite, are plausible enough. Axioms I_{14} and I_{15} could be replaced by some kind of Archimedean axioms. Moreover, the reader may find some relationship to Luce's <u>extensive</u> (<u>measurement</u>) <u>system</u>.

(ii) The axioms can be divided into three classes: First, those which point out the properties of $\parallel$; secondly, the axioms for $\preccurlyeq^{\circ}$; and thirdly, the interacting axioms giving the relationship between $\parallel$ and $\preccurlyeq^{\circ}$. There is no doubt about their consistency.

(iii) Instead of taking a Boolean algebra $\mathfrak{A}$, we could consider a complete complemented modular lattice, in which the relation $\perp$ would become a new primitive notion. In this case our axioms for $\perp$ and $\preccurlyeq^{\circ}$ come rather close to dimension theory of continuous geometry.

It is easy to show that Definition 8 implies Theorem 8, if we put $A \prec B \Longleftrightarrow B \preccurlyeq^{\circ} A \quad (A, B \in \mathfrak{A})$.

For purposes of representation we shall need a couple of notions which will be developed in the sequel.

Let $< \Omega, \mathfrak{A}, \preccurlyeq^{\circ}, \parallel >$ be a QI-structure. Then $\mathfrak{A}/\overset{\circ}{\sim} = \{[A]_{\overset{\circ}{\sim}} : A \in \mathfrak{A}\}$, where $[A]_{\overset{\circ}{\sim}} = \{B : A \overset{\circ}{\sim} B\}$. For simplicity we put $[A] = [A]_{\overset{\circ}{\sim}}$. Now we define a couple of operations on $\mathfrak{A}/\overset{\circ}{\sim}$:

(a) $[A] + [B] = [A_1 \cup B_1]$, if $A_1 \perp B_1$ and $A_1 \overset{\circ}{\sim} A \;\&\; B_1 \overset{\circ}{\sim} B$;

(b) $n \cdot [A] = (n-1) \cdot [A] + [A]$, $0 \cdot [A] = [\emptyset]$;

(c) $[A] \cdot [B] = [A_1 \cap B_1]$, if $A_1 \parallel B_1$ and $A_1 \tilde{\sim} A$ & $B_1 \tilde{\sim} B$;

(d) $[A]^n = [A]^{n-1} \cdot [A]$, $[A]^0 = [\Omega]$.

Axioms I_{12} and I_{13} will guarantee the correctness of the above definitions, that is, that they do not depend on the particular choice of representatives A_1, B_1. The existence of the defined terms is implied by I_{14} and I_{15}. Weakening of the axioms I_{14} and I_{15} would allow us to define only partial operations $+$, $n \cdot (-)$, $\cdot$, $(-)^n$ on $\mathcal{Cl}/\tilde{\sim}$.

We put, as might be expected,

$[A] \leq [B] \Leftrightarrow B \not\succ A \quad (A, B \in \mathcal{Cl})$.

The reader can easily develop the algebra of the ordered semiring $\mathbb{R} = \langle \mathcal{Cl}/\tilde{\sim}, [\emptyset], [\Omega], +, \cdot, \leq \rangle$. In particular, he can show that the operations $\cdot$ and $+$ are commutative, associative, monotonic, distributive, and the zero and unit element act as usual. Obviously, theorems like

$m \cdot [A] \leq n \cdot [A] \Leftrightarrow m \leq n$, provided $[A] \neq [\emptyset]$;

$[A]^n \leq [A]^m \Leftrightarrow n \leq m$, provided $[A] \neq [\Omega]$;

$(m+n) \cdot [A] = m \cdot [A] + n \cdot [A]$;

$[A]^{(m+n)} = [A]^m \cdot [A]^n$, are also true.

Our Representation Theorem for QI-structures is based on the existence of a function $\varphi : \mathbb{R} \longrightarrow Re$ such that

(i) $[A] \leq [B] \Leftrightarrow \varphi([A]) \leq \varphi([B])$,

(ii) $\varphi([\emptyset]) = 0$,

(iii) $\varphi([\Omega]) = 1$,

(iv) $\varphi([A] + [B]) = \varphi([A]) + \varphi([B])$, if $A \perp B$;

(v) $\varphi([A] \cdot [B]) = \varphi([A]) \cdot \varphi([B])$, if $A \parallel B$.

There are several ways of showing the existence of $\varphi : \mathbb{R} \longrightarrow Re$. We prefer here to use the method of Dedekind cuts of rational numbers.

In fact, the sets $c_B = \{ \frac{m}{n} : m \cdot [U] \leq n \cdot [B] \}$ and $c_B^* = \{ \frac{m}{n} : [U]^m \leq [B]^n \}$ form a Dedekind cut for fixed $U \in \mathcal{U}$,

since

(a) $m \cdot [U] \leq n \cdot [B] \veebar n \cdot [B] < m \cdot [U]$ and

$[U]^m \leq [B]^n \veebar [B]^n < [U]^m$ by I_7 *)

(b) $\frac{m}{n} \in c_B \ \& \ \frac{p}{q} \in c_B \Rightarrow \frac{m}{n} < \frac{p}{q}$ and

$\frac{m}{n} \in c_B^* \ \& \ \frac{p}{q} \in c_B^* \Rightarrow \frac{m}{n} < \frac{p}{q}$ by transitivity.

(c) $c_B^* = \emptyset$, defines 0 and c_B^* = set of all rationals, defines $+\infty$.

The real number which is defined by the Dedekind cut c_A (c_A^*) will be denoted by $\#c_A$ ($\#c_A^*$) . We shall define two real-valued functions on $\mathbb{R}$ as follows:

---

*) $\veebar$ denotes the logical connective 'exclusive or'

(1) $\varphi_u([U]) = u$, where $0 < u < 1$,

$\varphi_u([A]) = u \cdot \#c_A$

(2) $\varphi^*_v([U]) = v$, where $1 < v < +\infty$,

$\varphi^*_v([A]) = v^{-\#c_A}$

In the following we shall omit the indices u and v in functions φ_u and φ^*_v .

Using the consequences of axioms $I_1 - I_{15}$, it is easy to show that the conditions (i) - (v) hold for φ and φ^* . In fact,

$\varphi([A]) \leq \varphi([B]) \Longleftrightarrow u \cdot \#c_A \leq u \cdot \#c_B \Longleftrightarrow \{\frac{m}{n} : m \cdot [U] \leq n \cdot [A]\} \subseteq$

$\subseteq \{\frac{m}{n} : m \cdot [U] \leq n \cdot [B]\} \Longleftrightarrow [A] \leq [B]$. Similarly things hold for φ^* . If $A \perp B$, then $\varphi([A]) + \varphi([B]) =$

$= u \cdot \#c_A + u \cdot \#c_B = u \cdot (\#c_A + \#c_B) = n \cdot \#(c_A + c_B) =$

$= u \cdot \#c_{A \cup B}$, and similarly for φ^* .

$\varphi([A]) = \varphi([A \cup \emptyset]) = \varphi([A]) + \varphi([\emptyset])$, since $\emptyset \perp A$.

Hence, $\varphi([\emptyset]) = 0$. Again, $\varphi^*([\emptyset]) = \varphi^*([A \cap \emptyset]) =$
$= \varphi^*([A]) \cdot \varphi^*([\emptyset]) = 0$, since $\emptyset \parallel A$. In view of $\varphi[\emptyset] < \varphi([\Omega])$, we can normalize both φ and φ^* by taking

$$\frac{\varphi([A])}{\varphi([\Omega])} \quad \text{and} \quad \frac{\varphi^*([A])}{\varphi^*([\Omega])} \quad .$$

Now the fact that $\varphi([A]) \leq \varphi([B]) \Longleftrightarrow \varphi^*([A]) \leq \varphi^*([B])$ implies the existence of a one-one mapping $\eta : [0, 1] \rightarrow [0, 1]$ such that $\varphi^* = \eta \circ \varphi$.

Since $[A] \cdot ([B] + [C]) = [A] \cdot [B] + [A] \cdot [C]$, we get $\varphi([A] \cdot ([B] + [C])) = \varphi([A] \cdot [B]) + \varphi([A] \cdot [C])$, and so also

$$\eta^{-1}(\varphi^*([A]) \cdot \varphi^*([B] + [C])) + \eta^{-1}(\varphi^*([A]) \cdot \varphi^*([B])) +$$
$$+ \eta^{-1}(\varphi^*([A]) \cdot \varphi^*([C])) .$$

For $A \sim \Omega$ we get

$$\eta^{-1}(\varphi^*([B]) + \varphi^*([C])) = \eta^{-1}(\varphi^*([B])) + \eta^{-1}(\varphi^*([C])) .$$

But this is the Cauchy functional equation for η^{-1} in the real interval $[0, 1]$. Using the standard method of solution of linear functional equations, we get $\eta^{-1}(\varphi^*([A])) = \alpha \cdot \varphi^*([A])$, where α is a real positive constant. The normalization of φ and φ^* gives finally $\varphi^*([A]) = \varphi([A])$ for all $[A] \in \mathcal{A}/\approx$. We can now prove

THEOREM 18 (Representation Theorem) <u>Let</u> $< \Omega, \mathcal{A}, \precsim, \| >$ <u>be a</u> QI-structure. <u>Then there exists a finitely additive probability measure</u> P <u>on</u> $\mathcal{A}$ <u>such that</u> $< \Omega, \mathcal{A}, P >$ <u>is a probability space, and</u>

(1) $A \precsim B \Longleftrightarrow I(A) \leq I(B)$;

(2) $A \| B \Longleftrightarrow I(A \cap B) + I(A) + I(B)$;

(3) $I(A) = - \log_2 P(A)$.

Proof: We put $P(A) = \varphi([A])$ for $A \in \mathcal{A}$. Then from the previous discussion of φ it is easy to see that (1) - (3) are satisfied.

Clearly all the axioms I_1 - I_{13} are necessary conditions for the existence of the information measure I. Axioms I_{14} and I_{15} are not necessary. We leave open the problem of formulating axioms both necessary and sufficient for the existence of the measure I.

Aware of the relatively complicated necessary and sufficient conditions for the existence of a probability measure in an infinite Boolean algebra $\mathfrak{A}$, the author will not go here into further details.

$I(A) = -\log_2 P(A)$ is called sometimes as <u>self-information of the event</u> A. The next (slightly more general) notion is the so-called <u>conditional self-information of event</u> A, given event B: $I(A/B) = -\log_2 P(A/B)$. A further generalization leads to the <u>conditional mutual information of events</u> A <u>and</u> B, <u>given event</u> C:

$$I(A:B/C) = \log_2 \frac{P(AB/C)}{P(A/C)\cdot P(B/C)} \quad .$$

Naturally, we would like to give representation theorems also for these more complicated measures.

In this last case, our basic structure would be the set of complicated entities $A:B/C$ $(A, B, C \in \mathfrak{A},\ \emptyset \prec C)$ and two binary relations $\perp\!\!\!\perp$ and $\precsim$ on this set of entities. In fact, it would be enough to consider the formulas $A_1:B_1/C_1 \precsim A_2:B_2/C$ and $A/C \perp\!\!\!\perp B/C$, since the remainder can be defined as follows:

$$A:B \preccurlyeq C:D \Longleftrightarrow A:B/\Omega \preccurlyeq C:D/\Omega \ ;$$

$$A \preccurlyeq B \Longleftrightarrow A:A \preccurlyeq B:B \ ;$$

$$A/B \preccurlyeq C/D \Longleftrightarrow A:A/B \preccurlyeq C:C/D \ ;$$

$$A \parallel B \Longleftrightarrow A/\Omega \parallel B/\Omega \ , \quad \text{where } A,\ B,\ C,\ D \in \mathfrak{A} \ .$$

Some of the properties of the *qualitative conditional mutual information relation* $\preccurlyeq$ are analogous to those of the *qualitative self-information relation*. For example,

$$(A_1:C_1/E_1 \preccurlyeq A_2:C_2/E_2 \ \&\ B_1:D_1/E_1 \preccurlyeq B_2:D_2/E_2) \Longrightarrow A_1B_1:C_1D_1/E_1 \preccurlyeq A_2B_2:C_2D_2/E_2 \ ,$$

$$\text{if } A_i/E_i \parallel B_i/E_i \ \&\ C_i/E_i \parallel D_i/E_i \ \&\ A_iB_i/E_i \parallel C_iD_i/E_i \ , \quad i = 1, 2 \ .$$

We do not intend to develop further details here, because of the rather complicated nature of these properties. Note that we have several notions interacting here: conditional events, the independence relation, and the mutual information relation. From the point of view of algebraic measurement theory the problem is to give measurability conditions for very complicated relations defined on the above-mentioned complex entities.

4. APPLICATIONS TO PROBABILITY LOGIC, AUTOMATA THEORY, AND MEASUREMENT STRUCTURES

4.1. *Qualitative Probability Logic*

In methodology of science, inductive logic, and in philosophy generally, it is customary to consider the *probability of statements* rather than the *probability of events*. But even in the field of applied probability theory we quite often appear to speak of probabilities of statements rather than of sets. For example, we talk about the probability that the 'random variable ξ is not greater than the random variable η,' instead of taking the probability of the set $\{\omega \in \Omega : \xi(\omega) \leq \eta(\omega)\}$. This case, indeed, is nothing to worry about, since the appropriate translation from statements into events is immediately obvious. The main problem comes in when we want to talk of the probability of a statement containing quantifiers. The standard probability space $\mathbb{A} = \langle \Omega, \mathfrak{A}, P \rangle$ takes care at best only of the countable cases, so that the logical operations $\exists x$, $\forall x$ are often not adequately represented by the σ-operations in $\mathfrak{A}$; especially, when x runs over an uncountable domain. Consequently, the problem arises of how to assign a reasonable probability to quantified statements. The basic idea, following Scott and Krauss [20], is quite simple. We turn the Boolean algebra $\mathfrak{A}$, given in $\mathbb{A}$, into a <u>complete</u> Boolean algebra by taking the quotient $\mathfrak{A}/\Delta_P$, modulo the σ-ideal Δ_P of sets of measure zero. Then arbitrary Boolean operations are admitted. In addition, P turns into a strictly positive measure on $\mathfrak{A}/\Delta_P$. Therefore, if we assign homomorphically to every first-order formula an element of $\mathfrak{A}/\Delta_P$, no trouble will arise from using any sort of quantification. This should be clear enough. But the trick is not so innocent! Since $\mathfrak{A}/\Delta_P$ satisfies the countable chain condition, all Boolean operations actually reduce to countable ones; therefore the quantified formulas will get probabilities regardless of whether they are defined on a countable domain. Clearly some big Boolean algebras may be needed. But then we may not be able to guarantee the existence of a probability measure! Probability with values in a non-Archimedian field still may exist, but then we are faced with a problem of interpretation. In the author's opinion, the problem can be solved by considering a qualitative probability structure $<\Omega, \mathcal{A}, \preccurlyeq>$ for which, eventually, we will be prepared to give up the validity of the representation theorem. In fact, the formula $A \preccurlyeq B$ for $A, B \in \mathcal{A}$ has a perfectly good meaning or content in the above-mentioned fields, be it representable by a probability measure in the sense of problem (P_1) or not. In particular, $\mathcal{A}$ can be arbitrarily big, if needed. What matters now is only an appropriate way of assigning Boolean elements to formulas.

For this purpose consider a first-order language $\mathcal{L} =$ $= < V, F, P, \neg, \vee, \&, \Rightarrow, \Leftrightarrow, \forall, \exists >$, where V denotes the set of variables $x, y, z, v, w, \ldots,$ F the set of functors, P the set of predicates, and the remaining symbols stand for logical connectives and quantifiers in the usual way. Simplifying the problem, without losing generality, we shall consider just one two-place functor $\varphi \in F$ and one binary predicate $\rho \in P$. We define recursively first-order formulas over $\mathcal{L}$ in the well-known way. If needed, we may include among the logical symbols also the identy predicate $=$. We shall introduce Boolean models as probabilistic intended interpretations of $\mathcal{L}$. The aim is here to replace the truth values of ordinary logic by values in $\mathcal{Ot}$; then a formula is valid if it has value Ω , and invalid if it has value $\emptyset$. The various 'truth values' are ordered by the qualitative probability relation $\precsim$ of the qualitative probability structure $\mathbb{A} = \langle \Omega, \mathcal{Ot}, \precsim \rangle$ which will be held fixed throughout this section.

A nonempty set S together with a mapping $\equiv : S \times S \longrightarrow \mathcal{Ot}$ is called a Boolean set ($\mathbb{A}$-set) if and only if for all $a, b, c \in S$

(i) $[a \equiv a] = \Omega$;

(ii) $[a \equiv b \rightarrow b \equiv a] = \Omega$; *)

(iii) $[a \equiv b \cap b \equiv c \rightarrow a \equiv c] = \Omega$, where $a \equiv b = \equiv(a,b)$.

We could think of several mappings $\equiv$ on S , and they would yield different Boolean identity relations on S . If there is no danger of confusion we shall use S to refer to the structure $\langle S, \equiv \rangle$, and $S, S_1, S_2, \ldots$ will be variables for Boolean sets. Hence, roughly speaking, a Boolean set is just an ordinary set in which the natural identity is considered in terms of a Boolean-valued logic.

---

*) If $A, B \in \mathcal{Ot}$, then $A \rightarrow B$ denotes $\bar{A} \cup B$. There should be no confusion with the mapping f from A into B: $f : A \longrightarrow B$.

If $\equiv$ denotes the strict equality $=$ and $\mathfrak{A}$ is a two-element Boolean algebra, then $< S, \equiv >$ is equal to S .

A mapping $R : S \times S \longrightarrow Re$ is called a <u>Boolean binary relation</u> ($\mathbf{A}$ -relation) iff for all $a, b, c, d \in S$

$$[(a \equiv c \cap b \equiv d) \rightarrow (aRb \rightarrow cRd)] = \Omega ,$$

where $aRb = R(a,b)$.

It should be clear how one could define more general relations.

A Boolean relation R , defined on a Boolean set S , forms a <u>Boolean relational structure</u> ($\mathbf{A}$ -structure) $< S, R >$.

A mapping $f : S \times S \rightarrow S$ is called a <u>Boolean binary operation</u> ($\mathbf{A}$ -operation) iff for all $a, b, c, d \in S$

$$[(a \equiv c \cap b \equiv d) \rightarrow f(a,b) \equiv f(c,d)] = \Omega .$$

It is immediately clear how one gives a definition of <u>Boolean functions</u>.

A Boolean set S , together with a Boolean relation R and a Boolean operation f on it, defines a <u>Boolean structure</u> $< S, R, f >$.

Now we are ready to interpret the language $\mathcal{L}$ in a Boolean structure $< S, R, f >$, and give a definition of the qualitative probability formula $\Phi_1 \preccurlyeq \Phi_2$, where Φ_1, Φ_2 are formulas of $\mathcal{L}$.

We give values to variables $x, y, z, \ldots$ of V in the Boolean set S ; φ will denote a Boolean operation f in S and ρ will denote a Boolean relation R on S . Having done this, we get a possible Boolean model $\mathscr{S} = < S, R, f >$ for $\mathcal{L}$.

If the values of x,y are $\dot{x}, \dot{y} \in S$, then the value of the term $\varphi\, x\, y$ is $f(\dot{x},\dot{y})$. It is obvious how to extend this definition recursively to all terms.

Now the valuation $[\![\ ]\!]$ *) of formulas of $\mathcal{L}$ on $\mathcal{S}$ into $\mathfrak{A}$ is defined recursively as follows:

(i) $[\![\rho\tau_1\tau_2]\!]_{\mathcal{S}} = \dot{\tau}_1 R\, \dot{\tau}_2$, in particular, $[\![\tau_1 = \tau_2]\!]_{\mathcal{S}} =$
$= \dot{\tau}_1 \equiv \dot{\tau}_2$;

(ii) $[\![\neg \Phi]\!]_{\mathcal{S}} = \overline{[\![\Phi]\!]_{\mathcal{S}}}$;

(iii) $[\![\Phi_1 \vee \Phi_2]\!]_{\mathcal{S}} = [\![\Phi_1]\!]_{\mathcal{S}} \cup [\![\Phi_2]\!]_{\mathcal{S}}$;

(iv) $[\![\forall x\, \Phi]\!]_{\mathcal{S}} = \bigcap_{a \in S} [\![\Phi(a)]\!]_{\mathcal{S}}$, if $\Phi(a) = [x|a]\Phi$;

where τ_1, τ_2 denote terms, Φ, Φ_1, Φ_2 formulas of $\mathcal{L}$, and $[x|a]\Phi$ is a substitution operation in the metalanguage of $\mathcal{L}$.

We can put

$$\Phi_1 \preccurlyeq_{\mathcal{S}} \Phi_2 \Longleftrightarrow [\![\Phi_1]\!]_{\mathcal{S}} \preccurlyeq [\![\Phi_2]\!]_{\mathcal{S}} ,$$

and interpret $\Phi_1 \preccurlyeq_{\mathcal{S}} \Phi_2$ as follows: formula Φ_1 in the model $\mathcal{S}$ is not more probable than formula Φ_2 .

Considering all possible valuations $[\![\ ]\!]_{\mathcal{S}}$ we may define

$$\Phi_1 \preccurlyeq \Phi_2 \Longleftrightarrow [\![\Phi_1]\!]_{\mathcal{S}} \preccurlyeq [\![\Phi_2]\!]_{\mathcal{S}} \quad \text{for all } \mathcal{S} ,$$

and obtain a <u>qualitative probability structure of first-order formulas</u> $< \mathbb{F} , \preccurlyeq >$, in which, hopefully, the mentioned methodological

*) This ingenious notation is due to Scott and Krauss [20].

problems of empirical sciences can be studied.

Sometimes we start with a first-order theory $\mathcal{T}$ and take the class of all its models $\mathbb{M}_{\mathcal{T}}$. Then clearly

$$\Phi \in \mathcal{T} \Rightarrow [\,\Phi\,]_{\mathcal{S}} \sim \Omega \quad \text{for all} \quad \mathcal{S} \in \mathbb{M}_{\mathcal{T}}.$$

Note that in a qualitative probability structure of formulas $< \mathbb{F}\ ,\ \precsim\ >$ we are given a priori a fixed structure $\mathbb{A} = \; = < \Omega, \mathcal{U}, \precsim >$; and in the case of $< \mathbb{F}, \prec_{\mathcal{S}} >$ two structures, $\mathbb{A}$ and $\mathcal{S}$. The choice of $\mathcal{S}$ is given by empirical interpretation, but it is not clear, on the basis of which criteria should we choose $\mathbb{A}$.

One way of answering this question would be to associate with R a random relation R^*, that is, a mapping $R^* : \Omega \longrightarrow \mathcal{P}(S \times S)$,*) for which

$$\forall_{a,b\in S}[\{\omega : aR^*_{\omega}b\} \in \mathcal{U}\ ].$$

The random relation R^* is a random variable which takes as possible values ordinary relations on S. Now the randomization may be dictated by the empirical interpretation. In particular, we may be forced to take a special Ω, and $\mathcal{U}$ will be given by the conditions of observation. The subtlety of the events we

---

*) If A is a set, then $\mathcal{P}(A)$ denotes the set of subsets of A.

can observe will motivate us to choose an appropriate algebra from the lattice of algebras over Ω, ordered by the finer-than relation: $\mathcal{U}_1 \subseteq \mathcal{U}_2$. Finally, the probability relation $\preccurlyeq$ is given by the random mechanism of R*. If the randomization of R is not possible, we have to choose $\mathbb{A}$ subjectively.

If $<\Omega, \mathcal{U}, \preccurlyeq>$ is a qualitative conditional probability structure, then we can define the qualitative conditional probability relation on formulas from $\mathbb{F}$ as

$$\Phi_1/\Phi_2 \preccurlyeq \Psi_1/\Psi_2 \quad \| \Phi_1 \| / \| \Phi_2 \| \preccurlyeq \| \Psi_1 \| / \| \Psi_2 \| .$$

If we proceed in the same way as above and take a semiordered qualitative (conditional) probability structure, we can define notions like acceptability, rejectability, and the like. If needed, we can remove the condition that $\mathcal{U}$ be a Boolean algebra, and consider $\mathcal{U}$ as a lattice.

We shall not develop any specific details of these notions here.

4.2. Basic Notions of Qualitative Automata Theory

In this section an application of qualitative probability structures to probabilistic automata theory will be presented.

Automata theory is considered as a part of abstract algebra. Deterministic automata theory is a very well developed discipline, whereas probabilistic automata theory is still at the beginning stage. An excellent review of the subject can be found in R. G. Bucharaev [59].

Probabilistic automata represent empirical discrete systems in which statistical disturbances (noise) or uncertainties have to be taken into account. It is assumed also that the system has two channels: the <u>output</u> and <u>transition</u> channels.

From a formal point of view, a probabilistic automaton is a many-sorted structure*) $< \Xi, \Theta, \Sigma, H >$, where Ξ, Θ, Σ are finite nonempty sets (the set of <u>inputs</u>, the set of <u>outputs</u>, and the set of (<u>internal</u>) <u>states</u>) and H is a conditional probability function assigning to each 'conditional event' $(0,s')/(e,s)$ (where $0 \in \Theta$, $e \in \Xi$, and $s, s' \in \Sigma$) the probability that the automaton transits to state s' and produces output 0, given that the automaton is in state s with input e.

From a purely conceptual point of view, instead of taking H to be a mapping as above, that is, $H : \Xi \times \Sigma \longrightarrow \mathcal{D}(\Theta \times \Sigma)$, where $\mathcal{D}(\Theta \times \Sigma)$ denotes the set of probabilistic distribution functions over $\Theta \times \Sigma$, we can consider H to be a more general sort of mapping. In particular, we call the automaton $< \Xi, \Theta, \Sigma, H >$ <u>Boolean</u> if $H : \Xi \times \Sigma \longrightarrow \mathcal{U}^{\Theta \times \Sigma}$, where $\mathcal{U}$ is a Boolean algebra.**) Then $H((0,s')/(e,s))$ = the Boolean (truth) value of the statement that the automaton transits to state s' and produces output o, given that it is in state s with input e. In the Boolean algebra $\mathcal{U}$ we can have a qualitative probability relation $\preccurlyeq$,

*) By a many-sorted structure we mean a structure which has several different domains (universes).

**) If A and B are sets, then A^B denotes the set of mappings from B into A.

and therefore we can consider the qualitative probability formula

$(O_1, s_1')/(e_1, s_1) \precsim (O_2, s_2')/(e_2, s_2)$ $(O_1, O_2 \in \Theta,$ $e_1, e_2 \in \Xi,$ $s_1, s_1', s_2, s_2' \in \Sigma)$ with the obvious interpretation. Since we would not want to bother about the meaning of the algebra $\mathfrak{A}$, we shall proceed in a more straightforward way, namely, by replacing the function H by a qualitative probability relation. For this purpose, we have to consider <u>input events</u> (take just the elements of $\mathcal{P}(\Xi)$) and <u>state events</u> (take the elements of $\mathcal{P}(\Sigma)$). More specifically,

$$\begin{gathered}\text{if } O_1, O_2 \in \mathcal{P}(\Theta),\ e_1, e_2 \in \Xi,\ S_1', S_2' \in \mathcal{P}(\Sigma),\\ s_1, s_2 \in \Sigma, \text{ then } (O_1, S_1')/(e_1, s_1) \precsim (O_2, S_2')/(e_2, s_2) \Longleftrightarrow\\ \text{the output event } O_1 \text{ and the state event } S_1' \text{ given}\\ \text{input } e_1 \text{ and state } s_1 \text{ are not more probable than the}\\ \text{output event } O_2 \text{ and the state event } S_2' \text{ given input } e_2\\ \text{and state } s_2\,.\end{gathered} \tag{4.1}$$

This is the intended interpretation which we shall deal with. First comes the definition

DEFINITION 9 <u>A many-sorted structure</u> $< \Xi, \Theta, \Sigma, \precsim >$ <u>is called a finite qualitative probabilistic automaton</u> (FQP-automaton) <u>if and only if the following conditions are satisfied for all variables running over appropriate sets as explained in</u> (4.1):

M_0 $\Xi, \Theta,$ <u>and</u> Σ <u>are finite nonempty sets</u> (input, output, and state sets); <u>and</u> $\precsim$ <u>is a binary relation on</u>

$\mathcal{P}(\Theta) \times \mathcal{P}(\Sigma) \times \Xi \times \Sigma$, *where the formula generated by* $\precsim$ *is written as in* (4.1);

M_1 $(\emptyset,\emptyset)/(e_1,s_1) \precnsim (\Theta,\Sigma)/(e_2,s_2)$;

M_2 $(\emptyset,\emptyset)/(e_1,s_1) \precsim (O,S')/(e_2,s_2)$;

M_3 $(O_1,S_1')/(e_1,s_1) \precsim (O_2,S_2')/(e_2,s_2) \vee (O_2,S_2')/(e_2,s_2) \precsim (O_1,S_1')/(e_1,s_1)$;

M_4 $\forall_{i < n}[(O_i,S_i)/(e_i,s_i) \precsim (Q_i,S_i')/(e_i^*,s_i^*)]$

$$(Q_n,S_n')/(e_n^*,s_n^*) \precsim (O_n,S_n)/(e_n,s_n) \ \text{\underline{if}} \ \sum_{i \leq n} (O_i,S_i)^\wedge / (e_i,s_i) =$$

$$= \sum_{i \leq n} (Q_i,S_i')^\wedge / (e_i^*,s_i^*) .$$

We have mentioned many times that the characteristic function occurring now in M_4, can always be eliminated. To be completely clear, we put $[(O,S)^\wedge/e_1,s_1](o,s) = 1$ iff $o \in O \,\&\, s \in S$, otherwise zero. After those experiences obtained from manipulations with probabilistic relational structures, we might suspect that this definition is just the 'qualitative version' of the standard definition of probabilistic automaton. In fact, the following theorem can be easily proved.

THEOREM 19 *Let* $< \Xi, \Theta, \Sigma, \precsim >$ *be a many-sorted structure, described by axiom* M_0 *in Definition* 9. *Then it is a* FQP-automaton if and only if *there is a function* $H : \Xi \times \Sigma \longrightarrow \mathcal{D}(\Theta \times \Sigma)$ *such that* $< \Xi, \Theta, \Sigma, H >$ *is a probabilistic automaton* (especially, $H((o,s')/(e,s))$ is non-negative and $\sum_{\substack{o \in \Theta \\ s' \in \Sigma}} H((o,s')/(e,s)) = 1)$, and $(o_1,s_1')/(e_1,s_1) \precsim (o_2,s_2')/(e_2,s_2) \Leftrightarrow H((o_1,s_1')/(e_1,s_1) \leq$

$\leq H((o_2,s_2')/(e_2,s_2))$.

Taking H in the probabilistic automaton $\mathfrak{C} = \langle \Xi, \Theta, \Sigma, H \rangle$ to be a special function, we obtain, amongst others, the following classes of automata:

(i) $\mathfrak{C}$ is called a Mealy-automaton iff $H((o,s')/(e,s) =$

$= H(o/(e,s)) \cdot H(s'/(e,s))$.

(ii) $\mathfrak{C}$ is called a Moore-automaton iff $H(o/(e,s,s')) = H(o/s')$.

(iii) $\mathfrak{C}$ is called a probabilistic automaton with random output and deterministic transition iff

$H(s'/(e,s)) = 1$, if $\exists_f[s' = f(e,s)]$, and zero otherwise.

(iv) $\mathfrak{C}$ is called a probabilistic automaton with random transition and deterministic output iff

$H(o/(e,s)) = 1$, if $\exists_f[o = f(e,s)]$, and zero otherwise.

A special case of the Moore-automaton is the Rabin-automaton $\langle \Xi, \Sigma, \Psi, H \rangle$ where $\Psi = \{s : s \in \Sigma \ \& \ g(s) = 1\}$, where g is a mapping from Σ to Θ .

The qualitative version of these automata is quite obvious. In the case of Mealy-automata we have to require that $o/(e,s) \perp\!\!\!\perp s'/(e,s)$; and the appropriate axioms can be stated easily by using the results of Section 2.6 on qualitative conditional probabilities. Similarly, the Moore-automaton is specified by the requirement $o/s' \perp\!\!\!\perp (e,s)/s'$.

Notions like subautomaton, isomorphism and homomorphism of automata, reduction of states, direct sum and tensor product of automata, are quite easily defined. Since we are not going to develop any specific theory about the properties and mutual relationships of those notions, we shall not give any further definitions. The notion of the event x realized by qualitative probabilistic automaton is also easy to define.

If somebody wants to study semiordered qualitative probabilistic automata, he is welcome to do so. All obvious combinations of these notions are hardly supported at the present time by any empirical problem. On the other hand, from a theoretical point of view, they represent a good source of mathematically interesting theories.

4.3. Probabilistic Measurement Structures

The notion of a relational structure is fundamental in most current empirical theories. Various ordering structures furnish the common idealization of a large number of mathematical, physical, behavioral, and other scientific conceptual structures in which the notion of a relation occurs. However, in numerous instances in which these relational structures are applied, the situation or the problem is rather over-idealized. This is evidently the case, for example, in measurement. If the relations are determined by experiment or observation, undoubtedly they must be supposed to depend on chance. In repeated experiments or observations (under fixed conditions) we do not get unvarying results, because of 'noise,' an unavoidable phenomenon with statistical structure. For instance, it is quite common to describe the measurement of weight of a given set of objects using an equal-arm balance system by a binary relational formula a R b (object a is less heavy that object b). This method is completely correct if the weight-difference of objects a and b is essentially greater than the friction in the balance system and the statistical disturbance factors. But in the case of precise measurement with relatively small weight-differences the relation R would not serve as an adequate notion for the measurement problem. In this case we cannot use any more the 'yes-no' answers given by a R b or b R a , for if we repeat the measurement act several times, we may get different results contradicting each other. The relation a R b would hold with certain probability, approximated by the relative frequency of occurrences of a R b . Therefore the relation R has to be replaced or interpreted as a random relation which takes as possible values the ordinary relations. But then the appropriate order-homomorphism of this (random) measurement structure into the structure of reals must be random, too. In physics, clearly enough, classical quantities have to be considered as random variables, if their magnitudes are small and the molecular or other fluctuations are taken into account.

In econometrics or in psychology, especially in preference and utility theory, it is a well-known fact that inconsistencies may occur in a subject's preference ordering. The reason for this is simply that we are unable to perceive all relevant characteristics of the objects on which the preference is defined. Here again the random or probabilistic relation is the appropriate notion.

A Boolean relational structure $< S, R >$ is called a *qualitative probabilistic relational structure over* $< \Omega, \mathfrak{A}, \preccurlyeq >$ iff there is a random relation R^* on Ω corresponding to R ; $\mathfrak{A}$ is the Boolean algebra over which $< S, R >$ is defined, and $\preccurlyeq$ is a qualitative probability relation on $\mathfrak{A}$. If we replace $\preccurlyeq$ by a probability measure P we get a *(numerical) probabilistic relational structure*.

Note that qualitative probabilistic relational structures are generalizations of ordinary relational structures. In fact, all theorems and definitions of algebraic measurement structures given, for example, in Suppes and Zinnes [58] have probabilistic counterparts. We shall take one example.

DEFINITION 10 *A qualitative probabilistic binary relational structure* $< S, R >$ *over* $< \Omega, \mathfrak{A}, \preccurlyeq >$ *is called a qualitative probabilistic semiorder* (QPS-structure) if and only if *the following axioms are valid for all* $x, y, z, w \in S$:

V_1 $[\![xRx]\!] \sim \emptyset$;

V_2 $[\![xRy \ \& \ zRw \Rightarrow (xRw \vee zRy)]\!] \sim \Omega$;

V_3 $[\![xRy \ \& \ yRz \Rightarrow (xRw \vee wRz)]\!] \sim \Omega$.

If $< S, R >$ is a QPS-structure, then

(1) $[\![xRy \ \& \ zRw]\!] \preccurlyeq [\![xRw \vee zRy]\!]$;

(2) $[\![xRy \ \& \ yRz]\!] \preccurlyeq [\![xRw \vee wRz]\!]$;

(3) $[\![\, xRy \ \& \ yRz \,]\!] \preccurlyeq [\![\, xRz \,]\!]$;

(4) $[\![\, xRy \,]\!] \preccurlyeq \overline{[\![\, yRx \,]\!]}$.

The proofs would be worked in Boolean logic and then V_2 and V_3 would be applied. In fact, the proof goes exactly the same way as in ordinary logic, so that there is no need to repeat it here.

Even the representation theorem goes through, if we rewrite its proof into Boolean terms:

THEOREM 20 *Let* $< S, R >$ *be a finite qualitative probabilistic structure over* $< \Omega, \mathfrak{A}, \preccurlyeq >$. *Then it is a* QPS-structure *if and only if there is a random function* $U : S \longrightarrow Ra$ *) *and a random variable* $\eta > 0$ *such that for all* $x, y \in S$:

$$[\![\, xRy \,]\!] \leftrightarrow [\![\, U(x) \geq U(y) + \eta \,]\!] \sim \Omega \ . \quad **)$$

The proof is analogous to the case of ordinary semiorder structures. Note that $[\![\, U(x) \geq U(y) + \eta \,]\!] \Leftarrow \{\omega \in \Omega : U_\omega(x) \geq$ $\geq U_\omega(y) + \delta_\omega\} \in \mathfrak{A}$.

As a consequence we get $[\![\, xRy \,]\!] \sim [\![\, U(x) \geq U(y) + \eta \,]\!]$ which turns into equality in $\mathfrak{A}/\!\sim$.

Choice theory also gets its probabilistic version along these lines. A probabilistic linear ordering structure $< S, R >$ is

---

*) Ra denotes the set of random real variables.

**) If $A, B \in \mathfrak{A}$, then $A \leftrightarrow B$ denotes $AB \cup \overline{A}\overline{B}$.

represented by a <u>probabilistic utility function</u> $U : S \longrightarrow Ra$, where

$$| xRy | \sim | U(x) \leq U(y) | \quad \text{for all} \quad x, y \in S \ .$$

The relationship between probabilistic and ordinary relational structures can be given nicely by the following commutative diagram:

$$\begin{tikzcd} \langle S, R \rangle \arrow[r, "U"] \arrow[d, "e"'] & \langle Ra, \leq \rangle \arrow[d, "E"] \\ \langle S, R_e \rangle \arrow[r, "u"] & \langle Re, \leq \rangle \end{tikzcd}$$

where for $x, y \in S$: $| xRy | \sim | U(x) \leq U(y) |$;
$xR_ey \Longleftrightarrow u(x) \leq u(y)$, and $EU(x) = u(x)$, $EU(y) = u(y)$,
$e(R) = R_e$.

Roughly speaking, the ordinary relational structures are the 'averages' of probabilistic relational structures.

In ranking theory the well-known special sorts of probabilistic transitivities (see J. Marschak [60]) assure, in the qualitative version, the following form:

Let $\langle S, R \rangle$ be a qualitative probabilistic relational structure over $\langle \Omega, \mathfrak{A}, \preccurlyeq \rangle$ and let $A \sim \bar{A}$ for some $A \in \mathfrak{A}$.

Then R is called

(i) <u>weakly transitive</u> iff $(A \preccurlyeq | xRy \ \& \ yRz | \Longrightarrow A \preccurlyeq | xRz |)$;

(ii) <u>moderately transitive</u> iff $(A \preccurlyeq | xRy \ \& \ yRz | \Longrightarrow$
$| xRy \ \& \ yRz | \preccurlyeq | xRz |$;

(iii) strongly transitive iff $(A \prec \| xRy \ \& \ yRz \| \longrightarrow$

$\| xRy \vee yRz \| \preccurlyeq \| xRz \|$;

where $x, y, z \in S$.

There are many interesting problems here which we cannot discuss in this work.

5. SUMMARY AND CONCLUSIONS

5.1. Concluding Remarks

The main contribution of this work is stated in 10 definitions and 20 theorems. We have been studying in detail and under various conditions the properties of two binary relations $\preccurlyeq$ and $\preccurlyeq^{\cdot}$; the first one on Boolean algebras, and the second one on lattices of partitions. The results are quite general and simple, especially in finite structures.

Our basic concern was to show that probability, entropy, and information measures can be studied successfully in the spirit of representational or algebraic measurement theory.

The method used here is based on the most general results of modern mathematics, which state a one-one correspondence among relations, cones in vector spaces and the classes of positive functionals.

The main problems, stated in Section 1.1, have been solved in sufficient detail. In particular, we followed Scott in discussing the complete answer for (P_1). Answers were obtained for (P_2) and (P_3) only in the finite case and in a special form.

As applications, we solved similar problems for entropy, information, and automata.

As side problems, we discussed several conditional entities like A/B, A/P, and $\mathcal{P}_1/\mathcal{P}_2$ in a set-theoretic framework. We studied also the basic properties of the independence relation $\perp\!\!\!\perp$, and quadratic measurement structures. Various applications in logic, methodology of science, and measurement theory were indicated.

We have experienced the difficulties of measurement problems in the nonlinear case. Yet, only the successful solution of such cases is likely to persuade anyone to the importance of algebraic measurement theory, a theory which at present is still in rather a poor state.

As noted in Section 1.1, several people have tried to develop <u>semantic information theory</u>. In the author's view, it can be very well reduced to the standard information theory, because the set of propositions, on which semantic information measures are defined, forms, under certain rather weak conditions, a Boolean algebra. We do not think that there is much of learning about information measures on propositions, before a satisfactory theory of probability on first-order languages is developed. Probabilities of quantified formulas may then give something new. Beyond that there is the prospect of studying entropies in first-order theories and, perhaps, of answering some of the methodological questions posed by empirical theories. But any such advances will have to be preceded by elucidation of the structure of the independence relation on the set of quantified formulas, the structure of the set of conditional formulas, and so on. It may be that a purely qualitative approach would be more fruitful to begin with. Concerning these problems, in this study only the elementary facts have been stated.

The probability relation $\preccurlyeq$ is usually associated with subjectivist interpretations. The author has tried to show that the interpretation is unimportant; what matters really are the measurement-theoretic properties of this relation. Because of this, various semiorder versions of this relation have been also studied.

5.2. Suggested Areas for Future Work, and Open Problems

In this work several important problems have been left open, and others emerged during the research.

In particular we have not given any answer to the problem of uniqueness of probability, entropy, and information measures. In problem (P_4) we were unable to prove the multiplication law for the conditional probability measure.

Our study is entirely algebraic; we have not tried to introduce any topological assumptions for the relations $\preccurlyeq$, $\dot{\preccurlyeq}$; yet it is reasonable to assume that the answers to problems (P_2), (P_3), and (P_4) in the infinite case will lean heavily on the topological properties of $\preccurlyeq$ in $\mathcal{U}$.

We have been studying the structures $<\Omega, \mathfrak{A}, \preccurlyeq>$ and $<\Omega, \mathbb{P}, \preccurlyeq^{\bullet}>$ intrinsically; no doubt, mutual relationships between these structures have also some importance in illuminating the empirical notions of a micro- and macro-structure. Thinking along these lines, we could consider the category of qualitative probability structures and study their basic algebraic properties externally.

The structures $<\Omega, \mathfrak{A}, \preccurlyeq, \parallel>$, $<\Omega, \mathfrak{A}, \preccurlyeq^{\circ}, \parallel>$, and $<\Omega, \mathbb{P}, \preccurlyeq^{\bullet}, \parallel>$ have not been studied enough. We do not know, for instance, the necessary and sufficient conditions for pairs $<\preccurlyeq, \parallel>$, $<\preccurlyeq^{\circ}, \parallel>$, and $<\preccurlyeq^{\bullet}, \parallel>$ in order to be able to find appropriate probability, information, and entropy measures, respectively.

Yet another question is to determine the conditions to be imposed on the structures $<\Omega, \mathfrak{A}, \preccurlyeq^{\circ}, \parallel>$ and $<\Omega, \mathbb{P}, \preccurlyeq, \parallel>$ to ensure that the representation by information I and entropy H have the more specific form:

$$A \prec^{\circ} B \Longleftrightarrow \varepsilon + I(A) \leq I(B), \quad 0 < \varepsilon < +\infty, \quad A, B \in \mathfrak{A},$$

$$\mathcal{P}_1 \prec^{\bullet} \mathcal{P}_2 \Longleftrightarrow \varepsilon + H(\mathcal{P}_1) \leq H(\mathcal{P}_2), \quad 0 < \varepsilon < +\infty,$$

$$\mathcal{P}_1, \mathcal{P}_2 \in \mathbb{P}.$$

This question is motivated by the problem that arises in algebraic measurement theory when, because of errors, we have limited distinguishability.

A further generalization of the problem occurs when the error, rather than being constant, is taken as a function $\mathcal{E}$ of the event A or experiment $\mathcal{P}$.

Another problem is to find those conditions that must be imposed on $\langle \Omega, \mathfrak{A}, \precsim, \| \rangle$ or $\langle \Omega, \mathbb{P}, \precsim^{\cdot}, \| \rangle$ for the probability occurring in the information or entropy measure to have a specific distribution (Bernoulli, Binomial, Gaussian, for instance). In this case we might hope that the measures will be unique up to some reasonable group of transformations; moreover, the qualitative way of proving theorems may be more straightforward.

We have not given too many details about quadratic (or, generally, nonlinear) measurement structures in physics. Yet, there are clear measurement problems connected with the representation of such quantities for which the π-theorem holds.

Some of the questions of probability logic, probabilistic automata theory, and probabilistic measurement theory appeared to be important and we hardly could touch them.

The author is clearly aware of the rather introductory character of this study to the vast field of open problems in the measurement-theoretic approach to the notions of probability, information theory, and methodology of science; he hopes that further results will be forthcoming.

REFERENCES

[1] Steinhaus, H. Les probabilités dénombrables et leur rapport á la theorie de la measure. Fundamenta Mathematicae, 1923, 4, 286-310.

[2] Kolmogorov, A. N. Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin. 1933.

[3] Mises, R. von. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Deuticke, Vienna: 1931.

[4] Wald, A. Die Widerspruchsfreiheit des Kollektivbegriffs der Wahrscheinlichkeitsrechnung. Ergebwisse eines Math. Kolloquiums, 1937.

[5] de Finetti, F. La prévision: ses lois logiques, ses sources subjectives. Ann. Inst. Poincare, 1937, 7, 1-68. English translation in H. E. Kyburg, Jr., and H. E. Smokler (Eds.), Studies in subjective probability. Wiley, New York: 1964, 93-158.

[6] Savage, L. J. The foundations of statistics. Wiley, New York: 1954.

[7] Carnap, R. Logical foundations of probability. University of Chicago Press, Chicago: 1950.

[8] Jeffreys, H. Theory of Probability, 2nd ed., Clarendon Press, Oxford: 1948.

[9] Braithwaite, R. B. Scientific Explanation. Cambridge Univ. Press, Cambridge: 1951.

[10] Kraft, C. H., Pratt, J. W. & Seidenberg, A. Intuitive Probability on Finite Sets. Ann. Math. Stat., 1959, 30, 408-419.

[11] Scott, D. Measurement Structures and Linear Inequalities. Jour. Math. Psych., 1964, 1, 233-247.

[12] Scott, D. Linear Inequalities and Measures on Boolean Algebras (unpublished paper 1964).

[13] Stelzer, J. H. Some results concerning subjective probability structures with semiorders. Techn. Rpt. 115, 1967, Inst. for Math. Studies in the Soc. Sci., Stanford University, Stanford.

[14] Koopman, B. O. The axioms and algebra of intuitive probability. Annals of Math., 1940, 41, 269-292.

[15] Shimony, A. Coherence and the axioms of confirmation. Jour. Symb. Logic, 1955, 20, 1-28.

[16] Suppes, P. Set-theoretical structures in science. Inst. for Math. Studies in the Soc. Sci., Stanford University, Stanford 1967.

[17] Villegas, C. On qualitative probability σ-algebras. Annals of Math. Stat., 1964, 35, 1787-1796.

[18] Łoś, J. Remarks on foundations of probability. Proc. Intern. Congress of Mathematicians, Stockholm: 1962. Pp. 225-229.

[19] Gaifman, H. Concerning measures on first-order calculi. Israel J. Math., 1964, 2, 1-18.

[20] Scott, D. & Krauss, P. Assigning probabilities to logical formulas. In J. Hintikka & P. Suppes (Eds.) Aspects of Inductive Logic, 1966. North-Holl. Publ. Co., Amsterdam, 219-264.

[21] Rabin, M. O. & Scott, D. Finite automata and their decision problems. IBM Journal of Research and Development, 1959, 3, 114-125.

[22] Starke, P. H. Theorie stochastischer Automaten I, II. Elektronische Informationsverarberitung und Kybernetik, 1965, 1, 5-32 & 71-98.

[23] Rényi, A. On a new axiomatic theory of probability. Acta Math. Acad. Sci. Hung., 1955, 6, 285-335.

[24] Rényi, A. On conditional probability spaces generated by a dimensionally ordered set of measures. Teoria Verojatuosti i Primenenie (in Russian), 1956, 1, 61-71.

[25] Császár, A. Sur la structure des espaces de probability conditionelle. Acta Mathematica Acad. Sci. Hung., 1955, 6, 337-361.

[26] Sikorski, R. On analogy between measures and homomorphisms. Ann. Soc. Pol. Math., 1950, 23, 1-20.

[27] Halmos, P. R. Lectures on Boolean Algebras. Van Nostrand, New York: 1963.

[28] Evans, H. P. & Kleene, S. C. A postulational basis for probability. The Amer. Math. Monthly, 1939, 46, 141-148.

[29] Copeland, A. H. A postulational characterization of statistics. In J. Neyman (Ed.) Proceedings of the Berkeley Symp. on Math. Stat. & Prob., Univ. of California Press, California 1949, 51-61.

[30] Mazur, S. & Orlica, W. Sur les espaces metriques lineares, II. _Studia Math_. 1953, _13_, 137-179.

[31] Ptak, V. On a theorem of Mazur and Orlicz. _Studia Math_. 1956, _15_, 365-366.

[32] Luce, R. D. Semiorders and a theory of utility discrimination. _Econometrica_, 1956, _24_, 178-191.

[33] Scott, D. & Suppes, P. Foundational aspects of theories of measurement. _Journal of Symbolic Logic_, 1958, _23_, 113-128.

[34] Luce, R. D. & Tukey, J. W. Simultaneous Conjoint Measurement: A new type of fundamental measurement. _Jour. Math. Psych._, 1964, _1_, 1-27.

[35] Aczél, J., Pickert, G. & Radó, F. Nomogramme, Gewebe, und Quasigruppen. _Mathematica_ (Cluj), 1960, _2_(25), 5-24.

[36] Marczewski, E. A general scheme of the notions of independence in mathematics. _Bull. de l'Acad. Polonaise des Sci., Ser. Sci. math. et phys._, 1958, _6_, 731-736.

[37] Maeda, S. A lattice theoretic treatment of stochastic independence. _Jour. Sci. Hiroshima Univ._, 1963, _27_, Ser. A-1, 1-5.

[38] Varma, R. S. & Nath, P. Information Theory - A Survey. _Journal of Math. Sci._, 1967, _2_, 75-109.

[39] Fadeev, D. K. On the Concept of Entropy of a Finite Probability Scheme (in Russian). _Uspechi Mat. Nauk, New Ser._, 1956, _11_, 227-231.

[40] Khinchin, A. I. Der Begriffe der Entropien in der Wahrscheinlichkeitsrechnung. _Arbeiten zur Informationstheorie_ I. Berlin: Mathematische Forschungsberichte VEB, 1957.

[41] Tveberg, H. A new derivation of Information Functions. *Math. Scand.*, 1958, *6*, 297-298.

[42] Kendall, D. G. Functional Equations in Information Theory. *Zietschr. fur Wahr. und Verw. Geb.*, 1964, *2*, 225-229.

[43] Lee, P. M. On the Axioms of Information Theory. *Ann. Math. Stat.*, 1964, *55*, 415-418.

[44] Renyi, A. On Measures of Entropy and Information. In *Proc. 4th Berkeley Symp. Prob. Stat.*, V. I, University of California Press, pp. 457-561.

[45] Ingarden, R. S. & Urbanik, K. Information without Probability. Colloquium Mathematicum, 1962, *9*, 131-150.

[46] Ingarden, R. S. A simplified Axiomatic Definition of Information. *Bull. Acad. Sci., Ser. math., astr. et phys.*, 1963, *11*, 209-212.

[47] Ingarden, R. S. Simplified Axioms for Information without Probability. *Roczniki Polskiego Towarzystwa Matematycznego, Seria* I: Prace Matematyczne, 1965, *9*, 273-282.

[48] de Feriet, J. K. & Forte, B. Information et Probabilite. *C. R. Acad. Sci. Paris*, 1967, *265* A, 110-114.

[49] Kolmogorov, A. N. Three approaches to the Definition of the concept 'Quantity of Information.' *Problemy Peredaci Informacii* (in Russian), 1965, *1*, 3-11.

[50] Kolmogorov, A. N. Logical Basis for Information Theory and Probability Theory. IEEE *Transactions on Information Theory*, 1967, *IT-14*, 662-664.

[51] Belis, M. & Guiasu, S. A Quantitative-qualitative Measure of Information in Cybernetic Systems. IEEE Transaction on Information Theory, 1957, IT-4, 593-594.

[52] Weiss, P. Subjektive Unsicherheit und subjektive Information. Kybernetik, 1968, 5, 77-82.

[53] Raiffa, H., Pratt, J. W. & Schlaifer, R. The foundations of decision under uncertainty: an elementary exposition. Jour. Am. Stat. Assoc., 1964, 59, 353-375.

[54] Bar-Hillel, Y. & Carnap, R. Semantic Information. Brit. Journ. Phil. Sci., 1953, 4, 144-157.

[55] Hintikka, J. & Pietarinen, J. Semantic Information and Inductive Logic. In Hintikka, J. & Suppes, P. (Eds.) Aspects of Inductive Logic, Amsterdam: North-Holland Publ. Co., 1966, 96-112.

[56] Hintikka, J. The varieties of Information and Scientific Explanation. In Van Rootsclaar & Staal (Eds.), Logic, Methodology and Philosophy of Sciences III, Amsterdam: North-Holland Publ. Co., 1968, 311-331.

[57] Erdos, P. On the distribution function of additive functions. Ann. Math., 1946, 46, 1-20.

[58] Suppes, P. & Zinnes, J. Basic Measurement Theory, in Luce, R. D., Bush, R. & Galanter, E. Handbook of Mathematical Psychology, I, New York: J. Wiley & Sons, 1963, pp. 1-76.

[59] Bucharaev, R. G. Theory of Probabilistic Automata. Kibernetika (in Russian), 1968, 2, 7-22.

[60] Marschak, J. Binary-choice constraints and random utility indicators. In K. J. Arrow, S. Karlin & P. Suppes (Eds.) Mathematical methods in the social sciences, 1959. Stanford: Stanford University Press, 1960. Pp. 312-329.

(Continued from inside front cover)

91 P. Suppes. Information processing and choice behavior. January 31, 1966.

92 G. Groen and R. C. Atkinson. Models for optimizing the learning process. February 11, 1966. (Psychol. Bulletin, 1966, 66, 309-320)

93 R. C. Atkinson and D. Hansen. Computer-assisted instruction in initial reading: Stanford project. March 17, 1966. (Reading Research Quarterly, 1966, 2, 5-25)

94 P. Suppes. Probabilistic inference and the concept of total evidence. March 23, 1966. (In J. Hintikka and P. Suppes (Eds.), Aspects of Inductive Logic. Amsterdam: North-Holland Publishing Co., 1966. Pp. 49-65.

95 P. Suppes. The axiomatic method in high-school mathematics. April 12, 1966. (The Role of Axiomatics and Problem Solving in Mathematics. The Conference Board of the Mathematical Sciences, Washington, D. C. Ginn and Co., 1966. Pp. 69-76.

96 R. C. Atkinson, J. W. Brelsford, and R. M. Shiffrin. Multi-process models for memory with applications to a continuous presentation task. April 13, 1966. (J. math. Psychol., 1967, 4, 277-300).

97 P. Suppes and E. Crothers. Some remarks on stimulus-response theories of language learning. June 12, 1966.

98 R. Bjork. All-or-none subprocesses in the learning of complex sequences. (J. math. Psychol., 1968, 1, 182-195).

99 E. Gammon. The statistical determination of linguistic units. July 1, 1966.

100 P. Suppes, L. Hyman, and M. Jerman. Linear structural models for response and latency performance in arithmetic. (In J. P. Hill (ed.), Minnesota Symposia on Child Psychology. Minneapolis, Minn.: 1967. Pp. 160-200).

101 J. L. Young. Effects of intervals between reinforcements and test trials in paired-associate learning. August 1, 1966.

102 H. A. Wilson. An investigation of linguistic unit size in memory processes. August 3, 1966.

103 J. T. Townsend. Choice behavior in a cued-recognition task. August 8, 1966.

104 W. H. Batchelder. A mathematical analysis of multi-level verbal learning. August 9, 1966.

105 H. A. Taylor. The observing response in a cued psychophysical task. August 10, 1966.

106 R. A. Bjork. Learning and short-term retention of paired associates in relation to specific sequences of interpresentation intervals. August 11, 1966.

107 R. C. Atkinson and R. M. Shiffrin. Some Two-process models for memory. September 30, 1966.

108 P. Suppes and C. Ihrke. Accelerated program in elementary-school mathematics--the third year. January 30, 1967.

109 P. Suppes and I. Rosenthal-Hill. Concept formation by kindergarten children in a card-sorting task. February 27, 1967.

110 R. C. Atkinson and R. M. Shiffrin. Human memory: a proposed system and its control processes. March 21, 1967.

111 Theodore S. Rodgers. Linguistic considerations in the design of the Stanford computer-based curriculum in initial reading. June 1, 1967.

112 Jack M. Knutson. Spelling drills using a computer-assisted instructional system. June 30, 1967.

113 R. C. Atkinson. Instruction in initial reading under computer control: the Stanford Project. July 14, 1967.

114 J. W. Brelsford, Jr. and R. C. Atkinson. Recall of paired-associates as a function of overt and covert rehearsal procedures. July 21, 1967.

115 J. H. Stelzer. Some results concerning subjective probability structures with semiorders. August 1, 1967

116 D. E. Rumelhart. The effects of interpresentation intervals on performance in a continuous paired-associate task. August 11, 1967.

117 E. J. Fishman, L. Keller, and R. E. Atkinson. Massed vs. distributed practice in computerized spelling drills. August 18, 1967.

118 G. J. Groen. An investigation of some counting algorithms for simple addition problems. August 21, 1967.

119 H. A. Wilson and R. C. Atkinson. Computer-based instruction in initial reading: a progress report on the Stanford Project. August 25, 1967.

120 F. S. Roberts and P. Suppes. Some problems in the geometry of visual perception. August 31, 1967. (Synthese, 1967, 17, 173-201)

121 D. Jamison. Bayesian decisions under total and partial ignorance. D. Jamison and J. Kozielecki. Subjective probabilities under total uncertainty. September 4, 1967.

122 R. C. Atkinson. Computerized instruction and the learning process. September 15, 1967.

123 W. K. Estes. Outline of a theory of punishment. October 1, 1967.

124 T. S. Rodgers. Measuring vocabulary difficulty: An analysis of item variables in learning Russian-English and Japanese-English vocabulary parts. December 18, 1967.

125 W. K. Estes. Reinforcement in human learning. December 20, 1967.

126 G. L. Wolford, D. L. Wessel, W. K. Estes. Further evidence concerning scanning and sampling assumptions of visual detection models. January 31, 1968.

127 R. C. Atkinson and R. M. Shiffrin. Some speculations on storage and retrieval processes in long-term memory. February 2, 1968.

128 John Holmgren. Visual detection with imperfect recognition. March 29, 1968.

129 Lucille B. Mlodnosky. The Frostig and the Bender Gestalt as predictors of reading achievement. April 12, 1968.

130 P. Suppes. Some theoretical models for mathematics learning. April 15, 1968. (Journal of Research and Development in Education, 1967, 1, 5-22)

131 G. M. Olson. Learning and retention in a continuous recognition task. May 15, 1968.

132 Ruth Norene Hartley. An investigation of list types and cues to facilitate initial reading vocabulary acquisition. May 29, 1968.

133 P. Suppes. Stimulus-response theory of finite automata. June 19, 1968.

134 N. Moler and P. Suppes. Quantifier-free axioms for constructive plane geometry. June 20, 1968. (In J. C. H. Gerretsen and F. Oort (Eds.), Compositio Mathematica. Vol. 20. Groningen, The Netherlands: Wolters-Noordhoff, 1968. Pp. 143-152.)

135 W. K. Estes and D. P. Horst. Latency as a function of number or response alternatives in paired-associate learning. July 1, 1968.

136 M. Schlag-Rey and P. Suppes. High-order dimensions in concept identification. July 2, 1968. (Psychom. Sci., 1968, 11, 141-142)

137 R. M. Shiffrin. Search and retrieval processes in long-term memory. August 15, 1968.

138 R. D. Freund, G. R. Loftus, and R.C. Atkinson. Applications of multiprocess models for memory to continuous recognition tasks. December 18, 1968.

139 R. C. Atkinson. Information delay in human learning. December 18, 1968.

140 R. C. Atkinson, J. E. Holmgren, and J. F. Juola. Processing time as influenced by the number of elements in the visual display. March 14, 1969.

141 P. Suppes, E. F. Loftus, M. Jerman. Problem-solving on a computer-based teletype. March 25, 1969.

142 P. Suppes and Mona Morningstar. Evaluation of three computer-assisted instruction programs. May 2, 1969.

143 P. Suppes. On the problems of using mathematics in the development of the social sciences. May 12, 1969.

144 Z. Domotor. Probabilistic relational structures and their applications. May 14, 1969.